

TRABAJO FIN DE GRADO

Grado en ingeniería mecánica

**DETECCIÓN DE PUNTOS CALIENTES EN
ESTRUCTURAS UTILIZANDO ONDAS GUIADAS**



Memoria y Anexos

Autor:	Jesús Romero Lozano
Director:	Luis Eduardo Mujica
Co-Director:	Magda Ruiz
Convocatoria:	Mayo 2019

Resumen

Las estructuras mecánicas acaban degradándose con el tiempo ya que están sometidas a esfuerzos constantes. A causa de ello, acaban apareciendo daños o defectos en la estructura. Algunos se pueden apreciar a simple vista, en cambio otros afectan a la estructura sin que nos demos cuenta degradando sus propiedades mecánicas. Por este motivo se deben hacer controles periódicamente para prevenir un posible colapso de la estructura, ya que, un colapso sin previo aviso podría suponer una catástrofe.

Un factor muy importante a tener en cuenta en una estructura es la temperatura. Un aumento de temperatura puede fatigar el material y en consecuencia que se desgaste con mayor velocidad o incluso que rompa. A veces, dicho aumento puede significar un aviso de que la estructura está a punto de colapsar. Así mismo, este proyecto tiene como objeto detectar puntos calientes (puntos donde la temperatura es más elevada) utilizando como estructura de estudio una placa fina de aluminio. Dichos puntos calientes representan un defecto y se han generado en zonas aleatorias de la placa utilizando una pistola de calor.

Existen varios métodos aplicados a la detección de defectos en estructuras. En el presente proyecto se ha utilizado un método basado en la propagación de ondas elásticas. Dicha propagación se realiza con la ayuda de varios transductores piezoeléctricos (PZT), utilizando uno como actuador y los demás como sensores. Las señales captadas son analizadas y procesadas para detectar cada uno de los distintos defectos generados en la placa a través de un Análisis de Componentes Principales (PCA) y los estadísticos T2 y Q mediante el software R. Para ello se aplican diferentes tipos de escalado y diferentes métodos de cálculo y se comparan las posibles combinaciones para finalmente determinar cuál de ellas es mejor.

Resum

Les estructures mecàniques acaben degradant-se amb el temps ja que estan sotmeses a esforços constants. A causa d'això, acaben apareixent danys o defectes en l'estructura. Alguns es poden apreciar a simple vista, en canvi altres afecten a l'estructura sense que ens adonem degradant les seves propietats mecàniques. Per aquest motiu s'han de fer controls periòdicament per prevenir un possible col·lapse de l'estructura, ja que, un col·lapse sense previ avís podria suposar una catàstrofe.

Un factor molt important a tenir en compte en una estructura és la temperatura. Un augment de temperatura pot fatigar el material i en conseqüència que es desgasti amb més velocitat o fins i tot que trenqui. De vegades, aquest augment pot significar un avís que l'estructura està a punt de col·lapsar. Així mateix, aquest projecte té com a objecte detectar punts calents (punts on la temperatura és més elevada) utilitzant com a estructura d'estudi una placa fina d'alumini. Aquests punts calents representen un defecte i s'han generat en zones aleatòries de la placa utilitzant una pistola de calor.

Hi ha diversos mètodes aplicats a la detecció de defectes en estructures. En el present projecte s'ha utilitzat un mètode basat en la propagació d'ones elàstiques. Aquesta propagació es realitza amb l'ajuda de diversos transductors piezoelèctrics (PZT), utilitzant un com actuator i els altres com a sensors. Els senyals captats són analitzats i processats per detectar i localitzar cada un dels diferents defectes generats a la placa mitjançant una Anàlisi de Components Principals (PCA) i els estadístics T2 i Q mitjançant el programari R. Per a això s'apliquen diferents tipus d'escalat i diferents mètodes de càlcul i es comparen les possibles combinacions per finalment determinar quina d'elles és millor.

Abstract

Mechanical structures end up degrading with time since they are subjected to constant efforts. As a result, damage or defects appear in the structure. Some can be perceived at first glance, while others affect the structure without us noticing, degrading its mechanical properties. For this reason controls should be made periodically to prevent a possible collapse of the structure, since a collapse without prior warning could be a catastrophe.

A very important factor to take into account in a structure is the temperature. An increase in temperature can fatigue the material and consequently wear more quickly or even break. Sometimes, such an increase may mean a warning that the structure is about to collapse. Likewise, this project aims to detect hot spots (points where the temperature is higher) using as a study structure a thin aluminum plate. These hot spots represent a defect and have been generated in random areas of the plate using a heat gun.

There are several methods applied to the detection of defects in structures. In the present project, a method based on the propagation of elastic waves has been used. This propagation is carried out with the help of several piezoelectric transducers (PZT), using one as an actuator and the others as sensors. Captured signals are analysed and processed to detect and locate each of the different defects generated on the plate through a Principal Components Analysis (PCA) and T2 and Q statistics using the R software. Different types of scaling and different calculation methods are applied and all the possible combinations are compared to finally determine which one is better.

Agradecimientos

Quiero agradecer a mi tutor de proyecto, el profesor Luis Eduardo Mujica, por ayudarme y enseñarme en todo lo que ha podido durante la realización del proyecto y, sobretodo, por su paciencia y comprensión.

Agradezco también a Magda Ruiz, codirectora del proyecto, su apoyo e implicación, así como sus consejos para redactar la memoria y pulir todos los detalles.

Agradezco al departamento de fluidos de la Escola d'Enginyeria de Barcelona Est (EEBE-UPC) por su colaboración habiendo prestado su cámara termográfica para la realización del proyecto.

Finalmente, quiero agradecer a mi familia y a mi pareja por darme fuerzas y ayudarme en los momentos difíciles.



Índice

Resumen.....	i
Resum.....	ii
Abstract.....	iii
Agradecimientos.....	v
1. Introducción.....	9
1.1. Motivación.....	10
1.2. Objetivos.....	12
2. Monitorización y detección de defectos en estructuras (SHM).....	14
2.1. Definición.....	14
2.2. Métodos.....	14
2.3. Ondas elásticas.....	15
2.4. Efectos de la temperatura en las ondas de Lamb.....	21
3. Montaje experimental.....	23
3.1. Infraestructura.....	23
3.2. Configuración de los experimentos.....	37
3.3. Parámetros.....	43
3.4. Experimentos realizados.....	46
3.5. Organización de los datos en R.....	49
4. Análisis de componentes principales (ACP).....	50
4.1. Conceptualización.....	50
4.1.2. Escalado o normalización de los datos.....	50
4.1.2. Covarianza y correlación.....	51
4.1.3. Transformación de la matriz de datos.....	52
4.1.4. Reducción de la dimensión de la matriz de datos transformada.....	55
4.1.5. Métodos de cálculo de la matriz de transformación lineal P.....	56
4.1.6. Ejemplo aplicación PCA.....	57
4.2. Índices de detección de anomalías.....	63
4.3. PCA aplicado al SHM.....	66
5. Análisis preliminar de datos.....	71

6. Detección y localización de puntos calientes.....	76
6.1. Resultados.....	76
6.1.1. Scores.....	77
6.1.2. Estadístico T^2	83
6.1.3. Estadístico Q.....	88
6.1.4. T^2 vs Q.....	93
7. Introducción a la localización de defectos.....	97
7.1. Metodología para la localización de defectos.....	98
7.2. Análisis preliminar de las contribuciones.....	99
8. Análisis de impacto ambiental.....	102
Conclusiones.....	103
Presupuesto y/o análisis económica.....	105
Lista de tablas.....	107
Lista de figuras.....	108
Bibliografía.....	113
Anexo A. Guía de manejo en R de los datos de los experimentos.....	116
A1. Preparación.....	116
A2. Selección de experimentos.....	118
A3. Representación gráfica.....	118
Anexo B. Guía de modelado de datos y realización de PCA.....	120
B1. Preparación.....	120
B2. Preproceso de datos.....	122
B3. Creación del modelo y cálculo del PCA.....	123
B4. Visualización de resultados de los scores.....	124
B5. Visualización de resultados para T2-statistic y Q-statistic.....	126
Anexo C. Introducción al cálculo de contribuciones con R.....	129
C1. Cálculo de Q por residuals y por loadings.....	129
C2. Cálculo y representación gráfica de las contribuciones de Q.....	130

1. Introducción

Desde siempre, detectar daños en estructuras cuando estos aún se encuentran en su fase inicial ha sido un reto para la ingeniería (ya sea mecánica, civil, aeronáutica, etc.). Para poder conseguirlo, se realizan estudios constantemente dentro del campo la monitorización y detección de daños en estructuras, conocida en inglés como SHM (Structural Health Monitoring). Ésta consiste en diseñar sistemas incorporados en la estructura que registran continuamente la actividad mecánica (Figura 1). Las investigaciones realizadas en los últimos años nos permiten desarrollar e implementar nuevos métodos y sistemas para la monitorización de las fallas en estados incipientes. Estos sistemas son no destructivos, es decir, permiten evaluar el daño en una estructura determinada sin alterar sus propiedades. Actualmente ya existen técnicas no destructivas para la localización de defectos.

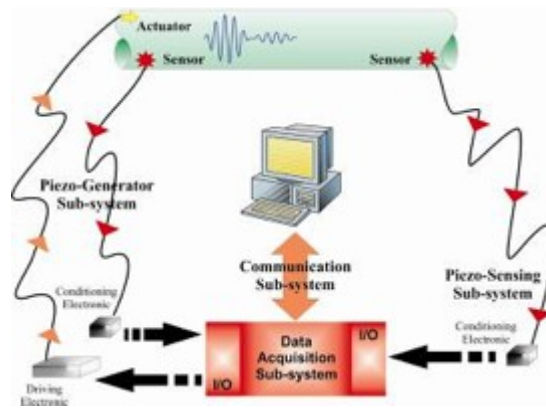


Figura 1.1: Monitorización y detección de daños en estructuras - SHM [1].

El grupo de investigación del laboratorio de Control, Dinámica y Aplicaciones (CoDALab) del departamento de matemáticas de la Universitat Politècnica de Catalunya (UPC) ha desarrollado y sigue trabajando activamente en técnicas para la evaluación de daños en estructuras (Figura 2). Estas se basan en la detección de variaciones y alteraciones en la propagación de ondas de Lamb utilizando los transductores que son elementos que transforman las características físicas y mecánicas en señales reconocibles. Estas señales pueden ser de varios tipos, como por ejemplo: de presión, de temperatura, de aceleración, de fuerza, entre otras.

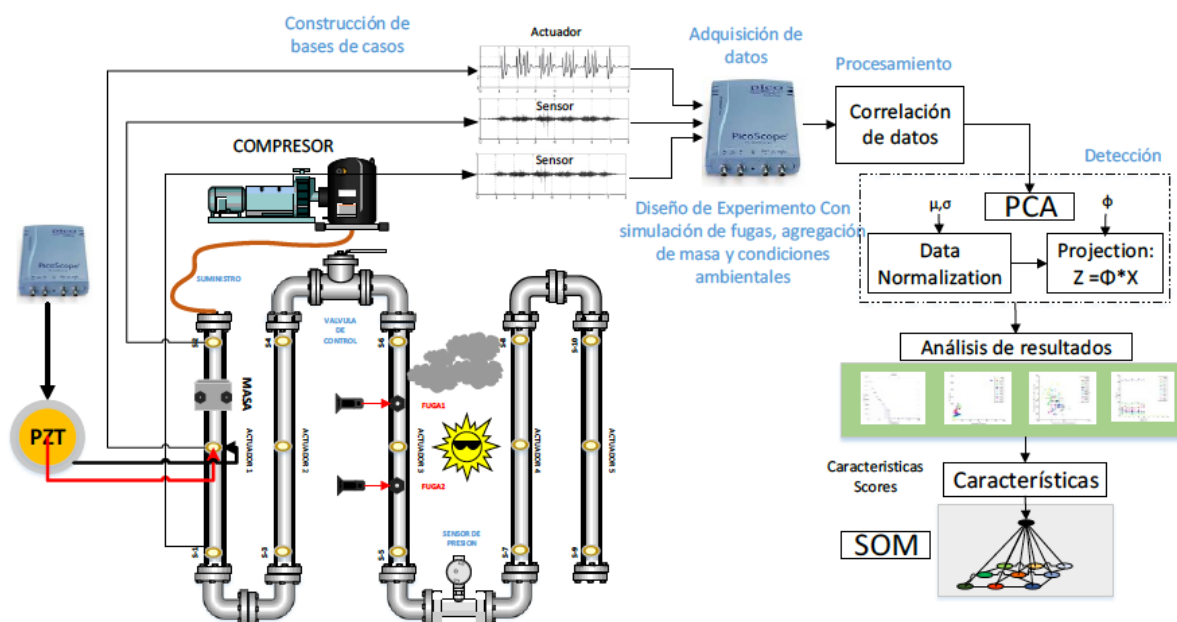


Figura 1.2: Metodología desarrollada en CoDALab para SHM [Propiedad del grupo CoDALab]

En SHM uno de los transductores más utilizados son los piezoeléctricos (PZT's), pueden ser instalados en las superficies de las estructuras y operar como sensores y actuadores. Los PZT's trabajan bajo el principio de la piezoelectricidad, que es un fenómeno físico que ocurre en determinados cristales cuando, al ser sometidos a tensiones mecánicas se genera una diferencia de potencial. Esto es debido a que los átomos del material se desplazan, sumando así sus cargas y creando dicha diferencia de potencial (voltaje). Este fenómeno también se da a la inversa, es decir, que es reversible. Gracias a los desplazamientos mecánicos de los transductores, las ondas de Lamb pueden propagarse más fácilmente ya que los PZT's a su paso se ven deformados y de nuevo transforman los desplazamientos mecánicos en señales eléctricas [2].

Detección de puntos calientes en estructuras utilizando ondas guiadas

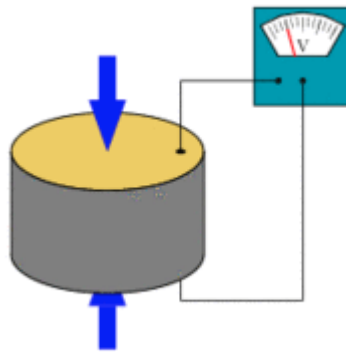


Figura 1.3: Principio de operación de los piezoelectricos (PZT's) [2].

Las ondas de Lamb son ondas acústicas que se propagan en placas sólidas. Son ondas elásticas cuyo movimiento de partículas se encuentra en el plano que contiene la dirección de propagación de la onda y la placa normal (la dirección perpendicular a la placa). Su comprensión y utilización ha aumentado considerablemente gracias al aumento de la potencia de cálculo. Por lo que se refiere a la estructura, si esta presenta un defecto, daño o deformación, la onda se altera y será diferente a la que se recibe en una estructura sana. Por esta razón los resultados obtenidos de un estudio de una estructura se deben comparar con los resultados que obtenemos de una estructura sana, para así llegar a conclusiones claras (Figura 1.4).

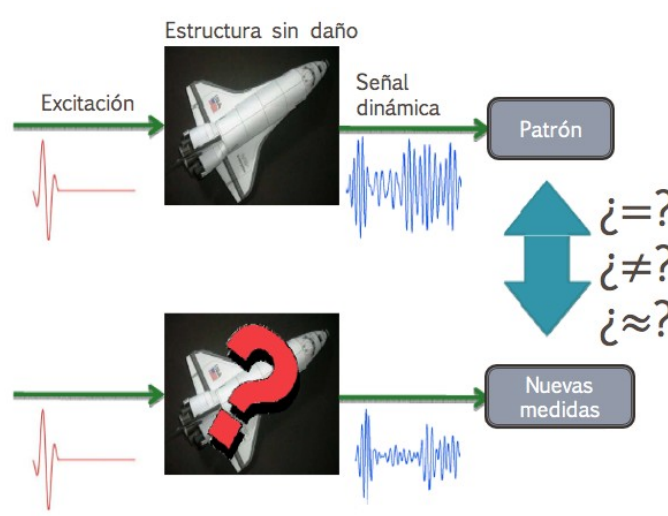


Figura1.4: Detección de daños o defectos en estructuras [Propiedad del grupo CoDALab].

1.1. Motivación

El equipo de investigación CoDALab lleva años enfrentándose a la detección de defectos estructurales. A través de estos años, se ha adquirido más experiencia y conocimiento. Gracias a ello, se ha avanzado en los procedimientos de captación, estudio de datos y análisis de resultados.

Los defectos en estructuras, pueden surgir a causa de muchos factores, uno de ellos puede ser la temperatura. El aumento de ésta en una región pequeña de la estructura, en muchas ocasiones puede tener graves consecuencias, así como fatigar el material, crear deformaciones irreversibles e incluso crear roturas que pueden afectar gravemente a la estructura y acabar en una catástrofe. El equipo CoDALab ya se había enfrentado antes a este factor, aunque siempre elevando la temperatura de toda la estructura y no calentando localmente. Por lo tanto, el cambio general de temperatura en la estructura se considera como una variación externa que afecta las señales adquiridas, pero en ningún caso representan defectos, de hecho, ocultan los defectos que pudiesen existir. Cambiando el enfoque, éste trabajo final de grado se centrará en la detección de puntos calientes en estructuras, es decir, en detectar si existe un gradiente de temperatura que pueda generar defectos en un futuro.

Para lograrlo, se calentarán puntos aleatorios en una placa de aluminio con la ayuda de una pistola de calor. Estos puntos representarán un defecto en la placa y a través de una cámara termográfica y con la ayuda de un termómetro con puntero láser podremos observar la temperatura a la que calentamos y observar si se mantiene constante a lo largo del experimento. Mediante ondas guiadas y con la ayuda de transductores piezoeléctricos se realizará la recogida de datos para su posterior estudio a través de un Análisis de Componentes Principales.

1.2. Objetivos

Un posible deterioro y defecto en la estructura se puede presentar cuando estas están siendo sometidas localmente a altas temperaturas por periodos prolongados de tiempo. Por otra parte, si cierta

Detección de puntos calientes en estructuras utilizando ondas guiadas

zona de la estructura presenta una alta carga, vibración, tensión, etc, su primer síntoma (antes de presentar defecto no reversible) es el calentamiento puntual.

Este tipo de defecto o anomalía aún no se ha montado en el laboratorio del grupo de investigación CoDALab. Por tanto, el objetivo principal y motivación de éste trabajo es la detección de puntos calientes en estructuras. Para lograrlo se emplearán las metodologías previamente trabajadas en CoDALab basadas en modelos estadísticos y se analizarán los diferentes resultados para detectar los puntos calientes en una estructura de laboratorio y así prevenir futuros desenlaces fatales.

Para poder lograrlo primero debemos establecer varios objetivos específicos:

1. Estudio y entendimiento de algunos conceptos de SHM y los diferentes tipos de sensores.
2. Obtener los datos por medio del montaje en laboratorio de un sistema completamente sensado de una estructura (placa de aluminio).
3. Aprender y entender los conceptos sobre los modelos estadísticos basados en el Análisis de las Componentes Principales ACP o mejor conocida por sus siglas en inglés PCA (Principal Component Analysis).
4. Aprender y establecer los algoritmos para este estudio estadístico mediante el empleo del software libre R con los datos almacenados en el laboratorio para así poder obtener conclusiones claras.

2. Monitorización y detección de defectos en estructuras (SHM)

2.1. Definición

El proceso de implementar una estrategia de identificación de daños o defectos para cualquier estructura dentro de la ingeniería (mecánica, civil, aeronáutica...) se conoce como *Structural Health Monitoring* (SHM). Se entiende como daño o defecto cualquier cambio en el material o en las propiedades de la estructura que afecten negativamente al rendimiento del sistema [3].

El objetivo principal de SHM es la evaluación continua, mediante ensayos no destructivos, del estado de las estructuras para garantizar una funcionalidad adecuada.

De entre los diferentes beneficios que presenta SHM, se puede destacar uno: nos permite conocer el comportamiento de la estructura bajo diferentes cargas y cambios ambientales y su estado actual para poder verificar si la estructura funcionará correctamente en un futuro o si debe ser reparada o reemplazada, con un bajo coste de mantenimiento.

2.2. Métodos

A lo largo de los años se han desarrollado varios métodos de SHM. En el siglo XIX, para evaluar si había daños en las ruedas de los ferrocarriles se utilizaba la vía acústica golpeando las ruedas con un martillo [3].

Durante décadas, se ha utilizado, y se sigue utilizando, una técnica basada en vibraciones. Esta depende de la vibración aplicada a través de sensores y ayuda a predecir la zona de riesgo en la estructura. Se debe tener cuidado con esta técnica, ya que la vibración puede dañar aún más a la estructura [4].

En el presente proyecto se utiliza un método basado en la propagación de ondas elásticas, en específico ondas de Lamb. Este tipo de método cada vez es más utilizado debido a que las ondas

Detección de puntos calientes en estructuras utilizando ondas guiadas

elásticas presentan ventajas como por ejemplo la velocidad de propagación en una distancia significativa o la alta sensibilidad a anomalías cerca de su trayectoria de propagación.

En general, un estudio basado en la identificación de daños en una estructura a partir de ondas elásticas, nos responde a tres preguntas que son claves para poder sacar conclusiones claras. Estas son:

1. ¿Hay daño en la estructura?
2. ¿Dónde está localizado?
3. ¿Es grave?

La primera pregunta nos da una percepción cualitativa de la presencia de daño. La segunda, nos muestra una localización cuantitativa. Finalmente, la tercera pregunta presenta información sobre la gravedad de ese defecto en la estructura.

2.3. Ondas elásticas

Una onda elástica la podemos definir como una perturbación que se propaga lo largo de un medio elástico (es decir, cualquier medio que tenga inercia y elasticidad).

A continuación se presentan diferentes tipos de ondas elásticas, junto con algunas de sus características:

Longitudinales

Las ondas longitudinales se desplazan a través del medio en la misma dirección de propagación de las ondas. Viajan como series de compresiones y rarefacciones alternantes (entendemos por rarefacción al proceso por el que un cuerpo se hace menos denso, es contrapuesto al fenómeno de compresión).

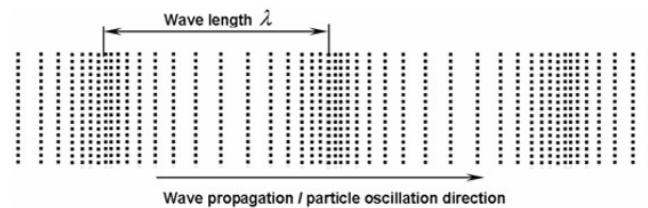


Figura 2.1: Representación esquemática de las ondas longitudinales [5].

Transversales

Son ondas que presentan oscilaciones en alguna dirección perpendicular a la dirección de propagación. Por ejemplo, si una onda se desplaza en el plano x-positivo, sus oscilaciones irán hacia arriba y abajo, es decir, en el plano y-z.

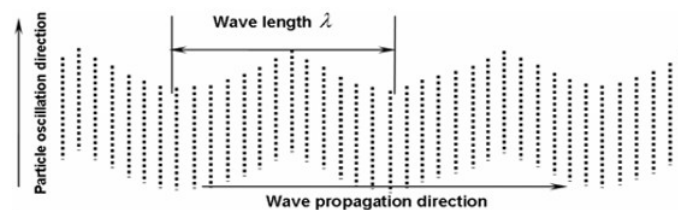


Figura 2.2: Representación esquemática de las ondas transversales [5].

Rayleigh

Estas ondas viajan a lo largo de las superficies de los sólidos. Se pueden producir por impacto localizado o por transducción piezoeléctrica. Estas ondas incluyen movimientos longitudinales y transversales que disminuyen exponencialmente en amplitud a medida que aumenta la distancia desde la superficie. Hay una diferencia de fase entre estos componentes del movimiento.

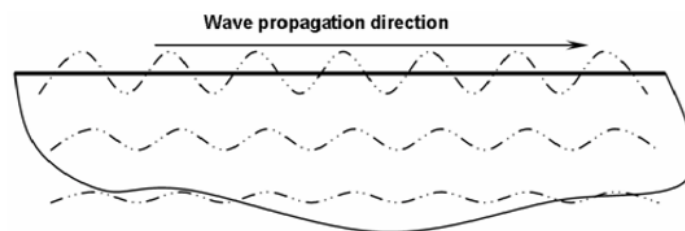


Figura 2.3: Representación esquemática de las ondas de Rayleigh [5].

Detección de puntos calientes en estructuras utilizando ondas guiadas

En sólidos isotrópicos, las ondas de Rayleigh hacen que las partículas de la superficie se desplacen en elipses en planos normales a la superficie y paralelos a la dirección de propagación.

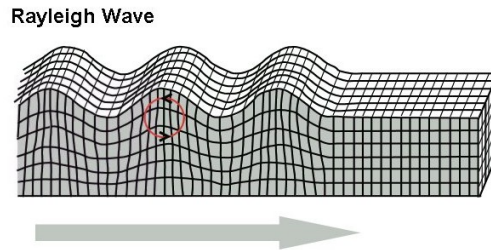


Figura 2.4: Representación esquemática de las ondas de Rayleigh en sólidos isotrópicos [6].

No son las únicas ondas superficiales que existen, hay otras como las ondas de Love o las ondas de Lamb (de las cuales hablaremos más adelante), ambas son ondas guiadas que viajan en masa.

La profundidad a la que podemos observar desplazamientos significativos en el sólido es aproximadamente igual a la longitud de onda.

Si comparamos las ondas de Rayleigh con las ondas transversales, se puede observar que las primeras tienen una velocidad ligeramente menor debido a un factor que depende de las constantes elásticas del material.

Ondas de Lamb

Las ondas de Lamb han encontrado una aplicación práctica sustancial, especialmente en el campo de los ensayos no destructivos. La aplicación de métodos que incluyen el uso de estas ondas hace que otros métodos para la detección de daños en estructuras sean sustituidos por éstos. Ya que el uso de las ondas de Lamb es menos costoso y más fácil de ser utilizado de forma continua.

Tal y como se ha comentado anteriormente, son ondas acústicas que se propagan en placas sólidas cuyo movimiento se encuentra en el plano que contiene la dirección de la onda y la dirección perpendicular a la placa. Sus propiedades son bastante complejas, ya que un medio infinito admite solo

dos modos de onda que viajan a velocidades únicas; pero las placas admiten dos conjuntos infinitos de modos de onda de Lamb, cuyas velocidades dependen de la relación entre la longitud de onda y el espesor de la placa.

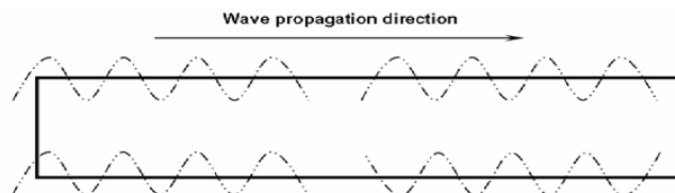


Figura 2.5: Representación esquemática de las ondas de Lamb [5].

Las ondas de Lamb manifiestan dispersión de velocidad, es decir, su velocidad de propagación c depende de la frecuencia (o longitud de onda) así como de las constantes elásticas y la densidad del material. El parámetro clave, es la relación entre el grosor de la placa d y la longitud de onda λ . Esta relación determina la rigidez efectiva de la placa y, por lo tanto, la velocidad de la onda.

$$f * d = \frac{d * c}{\lambda} \rightarrow c = f * \lambda \quad (1)$$

Dependiendo de la frecuencia, las ondas de Lamb tienen dos tipos de modos de vibración: modos de orden cero y modos de orden superior.

Modos de orden cero: tienen frecuencias nacientes de cero. Por esta razón, son los únicos modos que existen en todo el espectro de frecuencias, desde cero hasta frecuencias indefinidamente altas. Podemos distinguir dos subtipos de modos de orden cero: simétricos (S_0) o también denominados “modo extensional” y antisimétricos (A_0) o “modo de flexión”.

Los modos simétricos acostumbran a desplazarse de forma radial en el plano superficial de la placa, mientras que los modos antisimétricos acostumbran a desplazarse en la dirección del grosor de la placa.

Detección de puntos calientes en estructuras utilizando ondas guiadas

Estos modos llevan más energía que los de orden superior. El modo extensional a menudo aparece con mayor velocidad y menor amplitud que el modo de flexión.

Modos de orden superior: estos aparecen a medida que aumenta la frecuencia. Cada modo de orden superior nace a una frecuencia de resonancia de la placa y solo existen por encima de esa frecuencia.

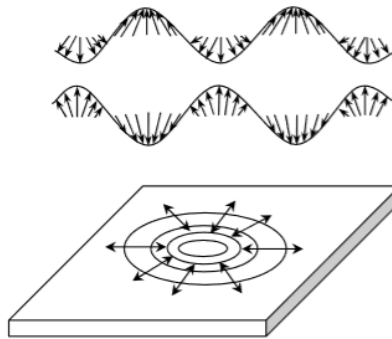


Figura 2.6: Comportamiento modo de vibración simétrico [5].

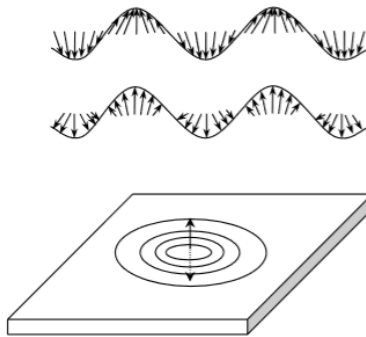


Figura 2.7: Comportamiento modo de vibración anti-simétrico [5].

Ondas de Stonely

Las ondas de Stonely normalmente se propagan a lo largo de una superficie sólido-sólido. También puede propagarse en superficies líquido-sólido, cuando esto ocurre se denominan ondas de Scholte.

La onda es de máxima intensidad en la superficie y disminuye exponencialmente lejos de ella.

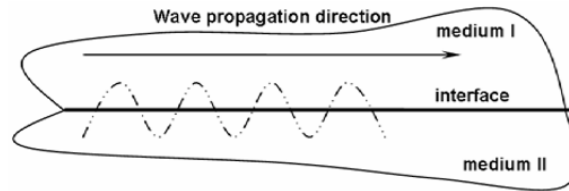


Figura 2.8: Representación esquemática de las ondas de Stonely [5].

Ondas de Love

Las ondas de Love son ondas superficiales polarizadas horizontalmente. Son el resultado de la interferencia de numerosas ondas transversales propagadas por una capa elástica superficial situada por encima de otro medio elástico. El movimiento de estas ondas se produce según una línea horizontal perpendicular a la dirección de propagación.

Moviéndose más profundamente en el material, el movimiento puede disminuir en un nodo y entonces aumentar y disminuir alternadamente cuando se examinan capas más profundas de partículas. La amplitud de dichas partículas disminuye rápidamente con la profundidad.

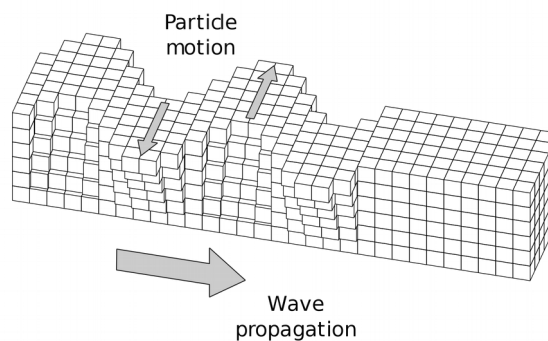


Figura 2.9: Representación esquemática de las ondas de Love [7].

2.4. Efectos de la temperatura en las ondas de Lamb

Los factores ambientales (temperatura, humedad...) pueden afectar tanto al funcionamiento y a las características de una estructura como a los análisis realizados para evaluar dicho funcionamiento. Depende de en qué condiciones se realice el análisis, los resultados obtenidos pueden variar.

En el caso de la temperatura, esta puede afectar en la propagación de ondas de Lamb sobre láminas finas. Existe una relación entre la temperatura y la magnitud de la amplitud de la señal: Cuando la temperatura aumenta, la amplitud de la señal también lo hace.

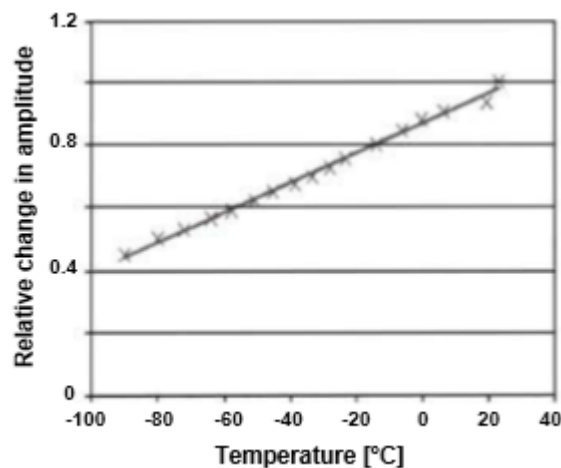


Figura 2.10: Efecto de la temperatura en la amplitud de la señal [5].

También se ha demostrado una relación entre la temperatura y el tiempo de propagación de la onda. Ocurre lo mismo que en el caso anterior, cuando la temperatura aumenta, el tiempo de propagación de la onda aumenta, lo que quiere decir que su velocidad disminuye.

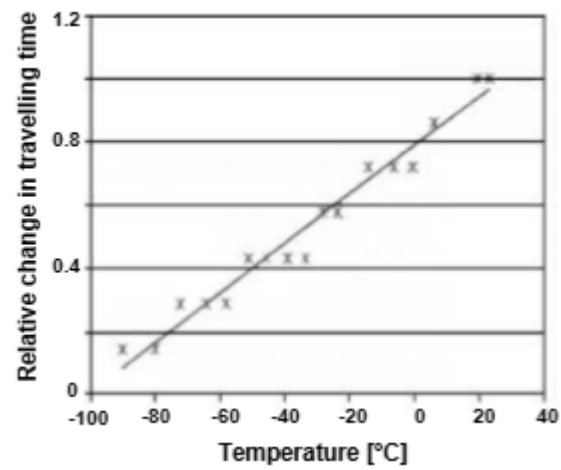


Figura 2.11: Efecto de la temperatura en el tiempo de propagación de la onda [5].

3. Montaje experimental

3.1. Infraestructura

Generador arbitrario de Onda

Para poder realizar el estudio deseado, es necesario poder generar una excitación/señal, el encargado de esta función es el generador arbitrario de onda. Esta señal será amplificada por un amplificador, del cual hablaremos más adelante y seguidamente se enviará a los actuadores piezoeléctricos. Para generar una onda determinada, es necesario definir unos valores para poder obtener la señal deseada. Una vez definidos serán transmitidos al generador arbitrario de onda. El modelo del que dispone el laboratorio del grupo CoDALab es un National Instruments PXI-5412.

A continuación encontramos algunas de las características del dispositivo:

- 1 canal: El generador dispone de un solo canal.
- 14 bits de resolución: Esto nos asegura un error mínimo debido a los valores que puede tomar la señal de salida. Ya que 14 bits de resolución implica que entre los valores extremos que podemos definir (es decir, entre el mínimo y el máximo) existen 16384 valores posibles para aproximar cada muestra.
- 100 MS/seg (millones de muestras por segundo): El generador es capaz de emitir señales tales que contengan hasta 100 millones de puntos por segundo.
- 20 MHz de frecuencia máxima para señales senoidales.
- 5.64 mV de amplitud en la escala mínima.
- 12 V de amplitud en la escala máxima.



Figura 3.1: Generador arbitrario de onda PXI-5412.

Digitalizador de señales

Este instrumento es el encargado de almacenar los datos de las variaciones de potencial (voltaje) que se generen en los sensores.

En el laboratorio del grupo CoDALab se dispone del modelo National Instruments PXI-5114. A continuación encontramos algunas de las características del dispositivo:

- 2 canales: Con un canal es suficiente para poder monitorizar daños en estructuras. El uso del segundo canal dependerá de si el estudio que se lleva a cabo requiere de señales adicionales.
- 8 bits de resolución: Esto implica que entre el mínimo valor que podemos digitalizar y el máximo existen 256 valores diferentes. Si comparamos los 14 bits de resolución comentados anteriormente del generador arbitrario de onda con los 8 bits del digitalizador de señales, se puede observar que en este caso el rango de puntos es bastante limitado. Por lo tanto es necesario definir la escala en la que digitalizamos las señales de modo que podamos obtener la resolución máxima (por ejemplo, no es

Detección de puntos calientes en estructuras utilizando ondas guiadas

adecuado digitalizar una señal que tiene por valor máximo 1 V en el rango de 10 V, ya que limitaríamos la resolución).

- 250 MS/seg (millones de muestras por segundo): El digitalizador de señales es capaz de digitalizar señales a una velocidad de muestreo de hasta 250 millones de puntos por segundo.
- 0.04 mV de amplitud en la escala mínima.
- 12 V de amplitud en la escala máxima.



Figura 3.2: Digitalizador de señales PXI-5114.

Conmutador de matriz y bloque terminal

El conmutador de matriz que podemos encontrar en el laboratorio CoDAlab es un modelo National Instruments PXI-2529. Por sí solo este dispositivo no tiene ninguna función, ya que requiere de un bloque terminal que contenga los terminales a los que conectaremos los cables y defina algún tipo de matriz (es decir, un conjunto de columnas y un conjunto de filas, de manera tal que cada columna se pueda conectar sólo a una fila). El bloque terminal del que dispone el laboratorio de CoDAlab es el modelo National Instruments TB2636. Éste define una matriz 4x32. Así pues, disponemos de 4 columnas que pueden ser conectadas a 32 filas (una columna sólo puede estar conectada con una fila simultáneamente). El generador arbitrario de onda amplificado se conectará a la columna 1, el digitalizador de señales a la columna 4 y los transductores piezoeléctricos a las filas, empezando por la fila 1 y hasta el número deseado.

Durante los experimentos, el conmutador irá conectando uno por uno los transductores que deban actuar como actuador al generador arbitrario de onda amplificado (columna 1) y los transductores que deban actuar como receptores al digitalizador de señales (columna 4).

A continuación encontramos algunas de las características del conmutador de matriz National Instruments PXI-2529 :

- Consta de 128 puntos de conexión.
- Soporta tensiones de hasta 150 V.
- Está basado en relés electromecánicos.
- Se puede configurar como dos tipos diferentes de matriz: 4x32 o bien 8x16.



Figura 3.3: Conmutador de matriz PXI-2529 y bloque terminal TB2636.

Chasis con controlador integrado

Los componentes comentados anteriormente no tienen forma alguna de comunicarse con el ordenador, por este motivo necesitamos el chasis con controlador integrado. El modelo del que disponemos es un National Instruments PXI-1033. Sus funciones son: proporcionar energía y la

Detección de puntos calientes en estructuras utilizando ondas guiadas

ventilación adecuada a todos los componentes mencionados en los puntos anteriores y hacer posible su comunicación con el ordenador.

Algunas características del dispositivo son:

- Velocidad de comunicación de hasta 110 MB/s.
- Acepta hasta 5 módulos.
- Suministro de potencia de 400 W.
- Controlador MXI-Express integrado.



Figura 3.4: Chasis PXI-1033.

Express card

Express card es un estándar de hardware que permite la conexión de otros dispositivos a un ordenador portátil. En nuestro caso servirá para conectar el chasis antes mencionado al ordenador (del cual hablaremos en el siguiente punto). El modelo de express card del que dispone el laboratorio CoDALab es un National Instruments Express Card-8360.



Figura 3.5: Express card NI-8360.

Ordenador

Para la utilización de PicoScope 5000 Series y la obtención de sus datos se necesita un ordenador capaz de soportar su software. En el laboratorio CoDAIab se dispone de un ordenador portátil modelo Dell Latitude E-6530.

Algunas de las características del ordenador las encontramos a continuación:

- Procesador Intel Core i5 3320M.
- 4 Gb de memoria RAM.
- Disco duro de 500 Gb.
- Windows 7 Professional.
- Slot para Express Card.
- Tres entradas USB: dos de ellas 2.0 y una 3.0.



Figura 3.6: Ordenador Dell Latitude E-6530 [8].

Amplificador de señal

Éste elemento es necesario ya que al tener un voltaje superior, la señal puede propagarse sobre la estructura a estudiar de manera más activa. El amplificador del que se dispone en el laboratorio es un modelo 7602M Wideband Amplifier de la marca Krohn-Hite.

Algunas de sus características principales son las siguientes:

- Potencia: 34watts a potencia continua, 68 watts a dc.
- Voltaje de salida de 282Vrms o 800Vpp.
- Voltaje de entrada de $\pm 200V$.
- Frecuencia de 1MHz.



Figura 3.7: Amplificador 7602M Wideband Amplifier [9].

Soldador

El soldador utilizado es el modelo VTSS4N. Se ha utilizado para realizar la conexión de los cables con los PZT's. Su rango de temperaturas es de 150°C a 450°C, la cual se regula mediante la rueda de selección de temperatura.



Figura 3.8: Soldador VTSS4N [10].

Transductores piezoeléctricos

Son dispositivos que utilizan el efecto piezoeléctrico (carga eléctrica acumulada en respuesta a la tensión mecánica y viceversa) para medir cambios de presión, aceleración, temperatura, tensión o fuerza al convertirlos en una carga eléctrica.

Su alto módulo de elasticidad los hace muy sensibles a la tensión mecánica. Esto le da al sensor una frecuencia natural extremadamente alta y una excelente linealidad en un amplio rango de frecuencias, haciéndolos adecuados para el acoplamiento ultrasonido. Además, son insensibles a los campos electromagnéticos y la radiación, lo que los hace adecuados en condiciones difíciles.

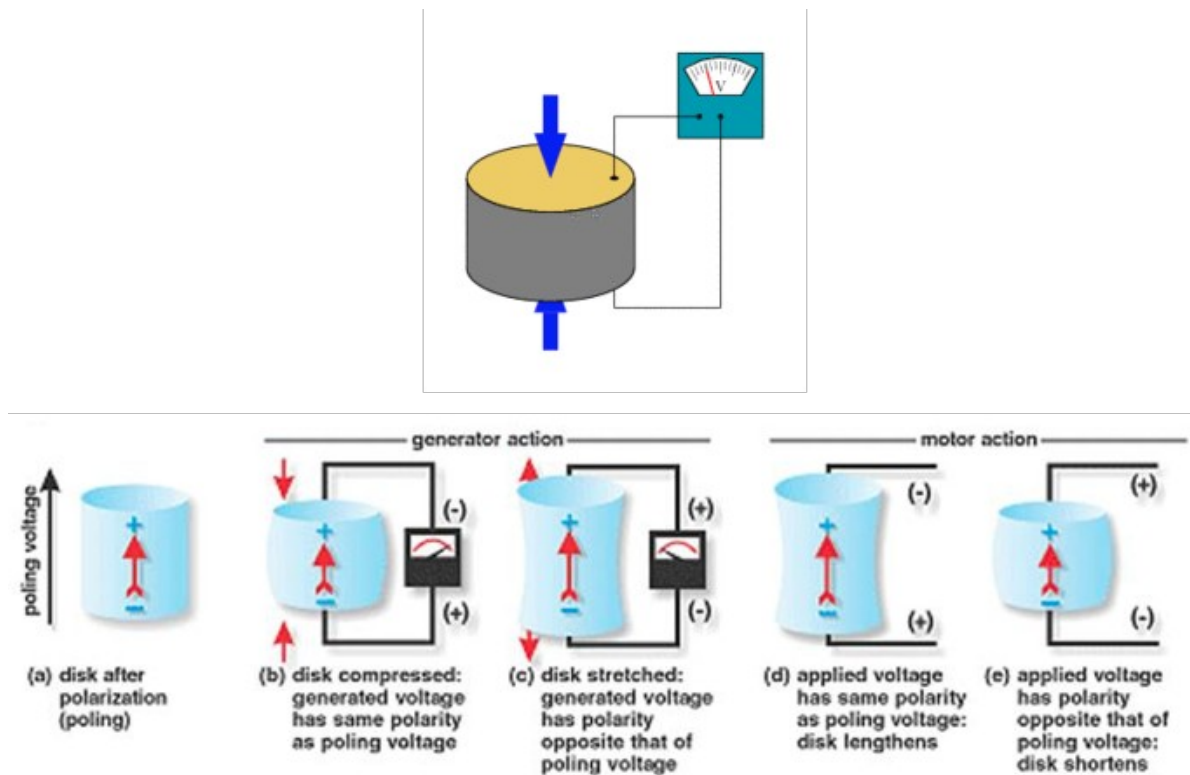


Figura 3.9: Funcionamiento de los piezoeléctricos [Propiedad del grupo CoDALab].

Pistola de aire caliente

La pistola de aire caliente se ha utilizado para aumentar la temperatura localmente en la estructura de estudio creando así zonas con temperatura más elevada. Estas zonas generadas son los denominados *Hot Spot* (punto caliente) y son los defectos a estudiar. La pistola utilizada es un modelo FME670K de la marca Stanley.

Estas son algunas de sus características:

- Potencia de entrada 2000W
- Temperatura de trabajo 50-450/70-600 °C
- Flujo de aire 300/550 litros/minuto



Figura 3.10: Pistola de aire caliente FME670K.

Cámara termográfica

Para verificar si realmente hemos calentado la estructura y para saber a qué temperaturas nos enfrentamos se ha hecho servir una cámara termográfica modelo PCE-TC 3 del fabricante PCE Ibérica. Ésta cámara ha sido proporcionada para realizar este proyecto por el departamento de fluidos de la *Escola d'Enginyeria de Barcelona Est* (EEBE), perteneciente a la *Universitat Politècnica de Catalunya* (UPC).

Algunas de sus especificaciones son:

- Rango de temperatura: $-10/+250$ °C.
- Tarjeta de memoria: hasta 1000 imágenes.
- Transferencia de datos para mediciones continuas a través de un puerto USB al PC.
- Tiempo operativo: 4-6 horas.
- Puntero láser: permite situar de forma exacta el campo de medición.



Figura 3.11: Cámara termográfica PCE-TC 3 [11].

Detección de puntos calientes en estructuras utilizando ondas guiadas

Para sujetar y mantener fija la posición de la cámara termográfica, se ha utilizado un trípode modelo 055PROB de la marca Manfrotto.

Termómetro con puntero láser

Este dispositivo se ha utilizado tanto para conocer la temperatura en la placa a la hora de calentar con la pistola de aire caliente FME670K, como para conocer a qué temperatura está la placa para poder realizar el siguiente experimento. El termómetro utilizado es un modelo IR608A de la marca AMPROBE.

A continuación se presentan algunas de sus especificaciones:

- Rango: $0^{\circ}F \rightarrow 750^{\circ}F / \square - 18^{\circ}C \rightarrow 400^{\circ}C$.
- Exactitud: $\pm 2\%$ de la lectura o $\pm 3,5^{\circ}F (\pm 2^{\circ}C)$ la opción que sea mayor.
- Tiempo de respuesta: 500ms.



Figura 3.12: Termómetro con puntero laser IR608A

Estructura de estudio

La estructura sobre la que se ha realizado el estudio y los experimentos del presente proyecto es una placa de aluminio de 1000 mm de ancho por 1500 mm de altura.



Figura 3.13: Estructura de estudio.

Osciloscopio

Este dispositivo no ha sido utilizado durante la configuración final de los experimentos, pero aún así es conveniente explicar sus características ya que se ha utilizado a lo largo del proyecto.

El laboratorio CodaLab dispone de un osciloscopio modelo PicoScope 5000 Series Flexible Resolution PicoScope de la casa Pico Technology. Éste nos permite construir las señales de excitación y registrar los datos recibidos por los sensores para así poder analizarlos. Se controla por ordenador a través de un software propio. Se conecta al ordenador mediante puertos USB, tiene una salida por

Detección de puntos calientes en estructuras utilizando ondas guiadas

donde se lanza la señal generada y cuatro entradas para poder adquirir los resultados de los experimentos.



Figura 3.14: PicoScope 5000 Series [12].

Para realizar las conexiones debemos tener algo en cuenta: Si se utilizaran un máximo de dos entradas el osciloscopio se conectará mediante los dos cables USB al ordenador. Por otra parte, si se utilizan más de dos entradas, se debe conectar el osciloscopio al ordenador mediante un cable USB y además, a la red eléctrica.

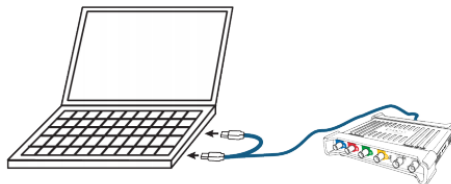


Figura 3.15: Diagrama de conexión máximo dos salidas [12].

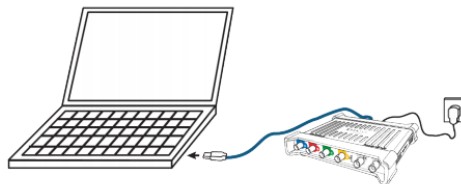


Figura 3.16: Diagrama de conexión mínimo tres salidas [12].

A continuación encontramos algunas características del osciloscopio PicoScope 5000 Series:

- Resolución de hardware seleccionable de 8 a 16 bits: Esto nos asegura un error mínimo debido a los valores que puede tomar la señal de salida. Ya que si seleccionamos la mínima resolución (8 bits) implica que entre los valores extremos que podemos definir (es decir, entre el mínimo y el máximo) existen 256 valores posibles para aproximar cada muestra. Mientras que si seleccionamos la resolución máxima (16 bits) existen 65536 valores posibles. Podemos observar que si seleccionamos la resolución mínima, el rango es bastante limitado. Por esa razón es necesario definir la escala en la que digitalizamos las señales de modo que podamos obtener la resolución máxima (por ejemplo, no es adecuado digitalizar una señal que tiene por valor máximo 1V en el rango de 10V, ya que limitaríamos la resolución).
- Rango voltaje máximo de medida: $\pm 20V$
- Protección sobretensiones: $\pm 100V$ en los canales de entrada y $\pm 20V$ en la señal del generador.
- Velocidad de muestreo: hasta 1GS/s
- Memoria de hasta 512 millones de muestras

Cada elemento del equipo tiene una función diferente, a continuación podemos encontrar donde se sitúa cada elemento y su respectiva función:

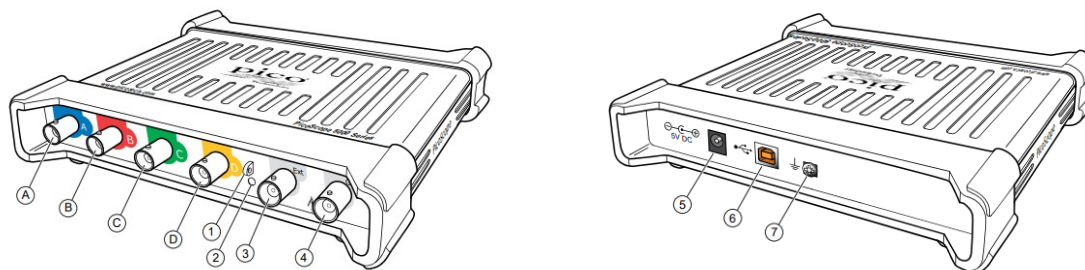


Figura 3.17: Diagrama de conexión PicoScope 5000 Series [12].

Detección de puntos calientes en estructuras utilizando ondas guiadas

A – Canal de entrada A

B – canal de entrada B

C – canal de entrada C

D – canal de entrada D

1 – Salida prueba compensación

2 – LED: De color rojo cuando el PicoScope está conectado pero no opera, de color verde cuando el osciloscopio captura datos.

3 – Entrada disparador externo

4 – Salida señas generador

5 – Toma de corriente DC

6 – Puerto USB 2.0

7 – Toma a tierra: Algunas baterías de portátiles pueden producir ruido. Para prevenir que este ruido interfiera en las medidas se puede conectar la toma a tierra a un punto a tierra externo. Alternativamente, se puede utilizar el ordenador portátil mientras se usa el cargador de batería.

3.2. Configuración de los experimentos

Configuración inicial

Inicialmente se iba a realizar un estudio sobre toda la superficie de la placa mediante 15 transductores piezoeléctricos que lleva pegados por una de sus caras. Para ello se utilizaría el chasis comentado anteriormente en el apartado de infraestructura junto con sus módulos correspondientes de los que dispone el laboratorio CoDALab.

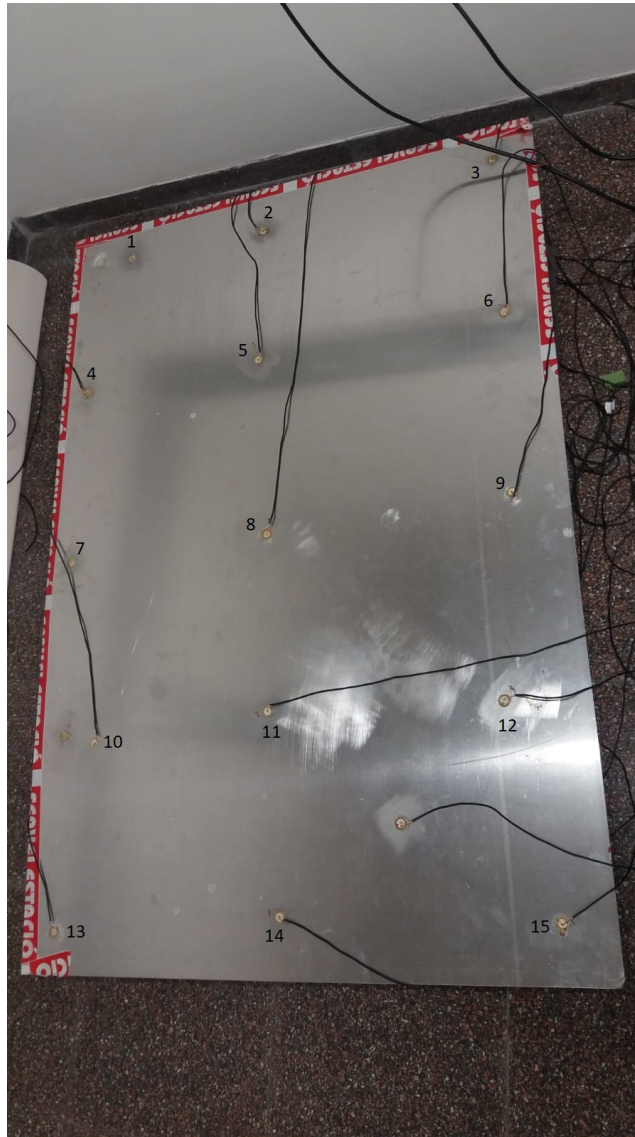


Figura 3.18: Configuración inicial: posicionamiento de los 15 PZT's.

Al empezar con la configuración y realización de los experimentos nos dimos cuenta de que no se obtenían resultados lógicos. Después de revisar todos los posibles fallos que podría haber, nos dimos cuenta de que el generador arbitrario de onda PXI-5412 estaba dañado por dos de sus conexiones y por lo tanto no era funcional. Se debía pensar una solución rápidamente para poder continuar con los experimentos. La solución pensada fue utilizar el PicoScope 5000 Series y, aunque mostraba algunas desventajas, era la única opción.

Detección de puntos calientes en estructuras utilizando ondas guiadas

Las desventajas principales que mostraba la realización de los experimentos con el PicoScope 5000 Series son: la primera, no podemos automatizar el proceso; la segunda, solo podemos utilizar 5 PZT a la vez (1 actuador y 4 sensores). Al tener que hacer manualmente todo el proceso de conexionado de los sensores para así conseguir todas las posibles combinaciones entre los PZT's con el fin de que todos actúen como sensor y actuador y además realizar manualmente el disparo de las ondas, se perdía demasiado tiempo.

Por esta razón se tuvo que tomar una segunda decisión importante: no se estudiaría toda la placa, sino solo una zona. En vez de los 15 PZT's iniciales, se decidió utilizar tan solo 5. Cuatro de ellos formando un cuadrado y en el centro de este está situado el quinto, el cual tiene la función de actuador. Al utilizar tan solo 5 PZT's, no hay que realizar cambios de conexionado y se pueden realizar los experimentos sin que la temperatura varíe mucho a lo largo de la duración de estos.

Una vez se realizaron los experimentos, se volvieron a observar datos muy diferentes a los que se deberían obtener, esto ocurría en todos los experimentos realizados. Vimos que siempre era el mismo canal el que nos daba unos datos anormales y se llegó a la conclusión de que un piezoeléctrico estaba dañado. Se estaba en lo cierto y efectivamente, al cambiarlo por uno nuevo los datos obtenidos eran normales.

Configuración final

Durante todo este proceso de configuración del PicoScope 5000 Series y experimentación, llegó la noticia de que el generador arbitrario de onda PXI-5412 estaba reparado. Así mismo se decidió volver a la idea original de utilizar los elementos de National Instruments descritos en el apartado de infraestructura aunque manteniendo la idea de utilizar tan solo 5 PZT. Al realizar los experimentos con la infraestructura de National Instruments, los datos se obtienen de una manera mucho más cómoda para poder realizar su estudio.

Se ha optado por utilizar los PZT situados en una de las esquinas de la placa para así ver diferentes comportamientos en los sensores. Si se hubiese seleccionado la zona central de la placa, existiría una simetría que no nos interesa, ya que se verían comportamientos muy similares.

En la siguiente imagen se puede apreciar la zona a estudiar de la placa así como la posición de los PZT's. La generación de los Hot Spots se ha realizado en la cara contraria a la que están pegados los PZT, esto se ha hecho para evitar tener cables de por medio que podrían suponer un estorbo.

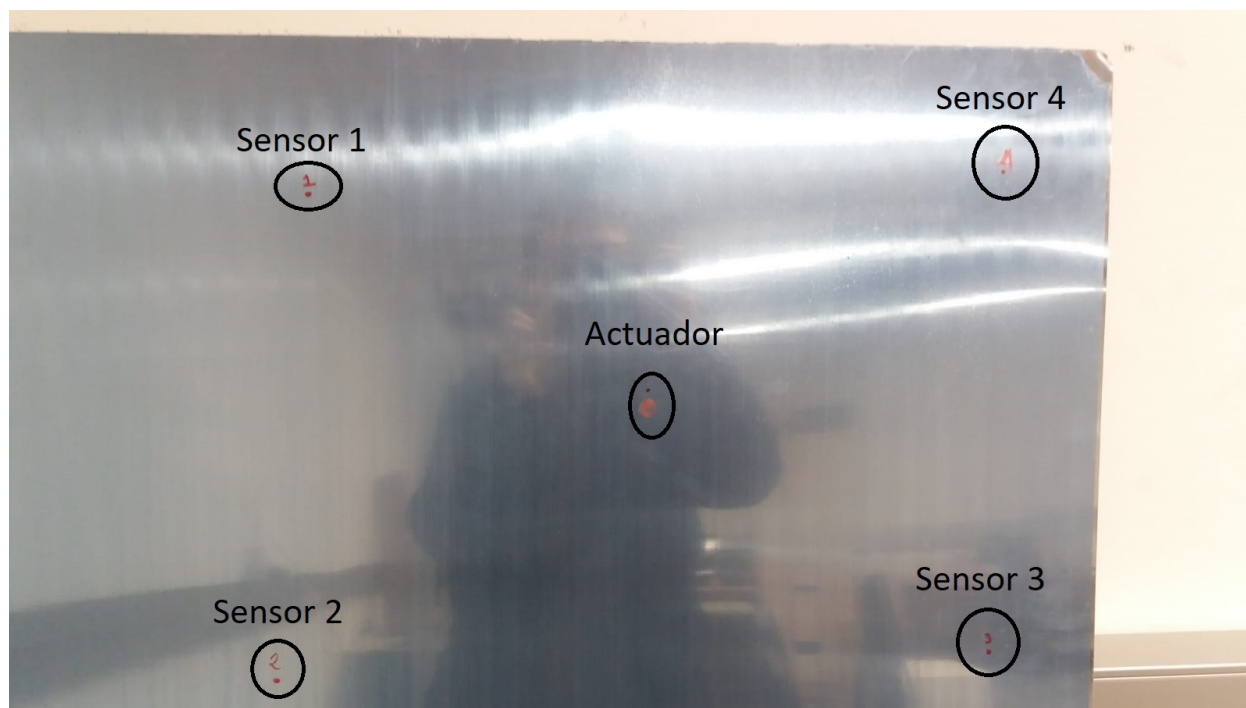


Figura 3.19: Zona de estudio y posicionamiento de los PZT.

Preparación de los experimentos

Para realizar el montaje de los transductores en la placa se han seguido una serie de pasos:

1. Limpieza de la zona de pegado del transductor piezoeléctrico

La zona en la que se posicionan los PZT's debe ser limpiada con tal de eliminar posibles impurezas en la superficie de la placa ya que pueden influir negativamente en la toma de datos. La limpieza se ha realizado con papel de lija.

Detección de puntos calientes en estructuras utilizando ondas guiadas

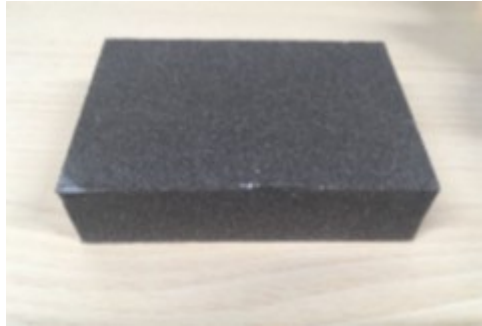


Figura 3.20: Papel de lija.

Una vez lijada la zona, se debe realizar una limpieza superficial con la ayuda de un trapo para eliminar los restos de polvo de la zona lijada.

2. Pegado de los transductores piezoeléctricos

Una vez limpiada la zona, se procede al pegado. Para ello se ha utilizado el pegamento *Araldite Cristal de dos componentes*. Es un pegamento de pegado rápido para todo tipo de materiales y al ser de dos componentes, es más resistente.



Figura 3.21: Pegamento Araldite Crystal.

Al usar el pegamento es aconsejable utilizar guantes de latex como protección para evitar dañarnos. Para aplicar el pegamento, se debe aplicar un poco del producto encima de una zona

acondicionada para ello (se puede utilizar por ejemplo un folio de papel) y se deben mezclar bien los dos componentes.

Una vez posicionado el PZT y con el pegamento aplicado, se debe mantener a una presión constante hasta que esté completamente pegado. Para ello se utiliza un peso pequeño, este se posiciona encima del PZT evitando dañarlo y evitando que toque el pegamento ya que podría quedar pegado a la estructura.

3. Soldado de los cables a los transductores piezoeléctricos

Una vez pegados los PZT, se deben conectar los cables para poder obtener la información. Nos encontramos con dos conexiones para cada cable, una roja y otra blanca. Como los PZT's utilizados tienen dos zonas polarizadas, cada conexión se debe soldar a una de ellas: el cable rojo a la zona interior y el cable blanco a la zona exterior.

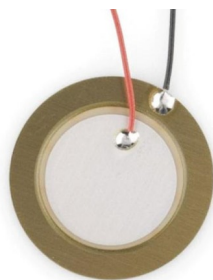


Figura 3.22: Conexionado PZT.

Respecto al montaje de la cámara termográfica, se ha utilizado un trípode para mantenerla fija apuntando hacia la placa. En un principio la placa estaba situada horizontalmente en el suelo, pero al fijarnos en la pantalla de la cámara observamos algún que otro problema. La pantalla es muy pequeña y para enfocar toda la zona de trabajo debía estar muy alejada de la placa y era un problema ya que el trípode no se podía elevar más. Además, en la placa se reflejaba el calor de los focos fluorescentes que iluminan el laboratorio. Por estas razones se decidió apoyar la placa en la pared y mantenerla de forma vertical.

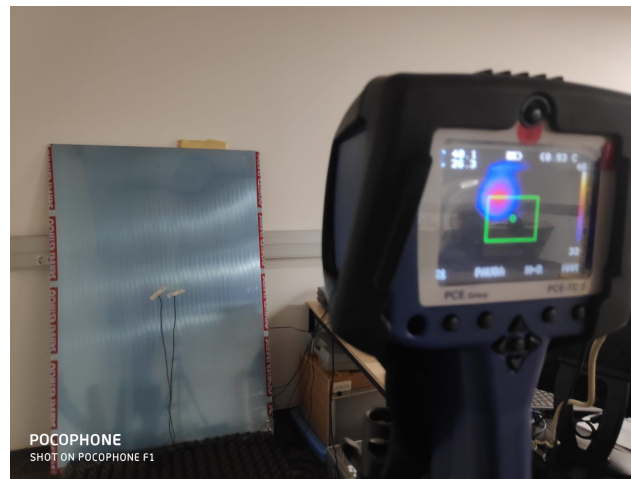


Figura 3.23: Posicionamiento de la cámara termográfica respecto a la placa.

3.3. Parámetros

Temperatura de calentamiento

Para la generación de los puntos calientes que representan los defectos sobre la estructura, se ha hecho uso de la pistola de calor FME670K. Para generar dicho defecto, se ha calentado siempre hasta alcanzar 70°C en el punto deseado para así tener la misma temperatura de referencia en todos los experimentos.

Para poder realizar el siguiente experimento, se ha esperado a que el punto caliente anterior se enfriase hasta llegar a 25-27°C. Para controlar cuándo se debía dejar de calentar dicho punto o cuando se podía empezar a realizar el siguiente experimento, se ha utilizado el termómetro con puntero láser IR608A, este nos da la temperatura al instante y podemos controlarla a la vez que calentamos.



Figura 3.24: Generación de puntos calientes y control de su temperatura.

Rango de temperaturas de la cámara

Para cada experimento se ha realizado una captura de imagen de la pantalla de la cámara termográfica. Para obtener en todos la misma gama de colores se decidió dejar un rango de temperaturas fijo: temperatura máxima de 70°C y temperatura mínima de 39°C.

Labview

El equipo CodaLab dispone de una interfaz para la detección de daños en estructuras la cual ha sido utilizada para la toma y el análisis de datos en el presente proyecto. A continuación se presentan los valores que se han utilizado para la configuración de los experimentos en la plataforma Labview:

- Número de actuadores y número de sensores: 1 actuador y 4 sensores
- Frecuencia base: un valor de 200 kHz. Nos indica la frecuencia inicial a la que se emitirá la señal.
- Voltaje base: un valor de 1V. Define la tensión de base a utilizar durante la tanda de experimentos.
- Número de repeticiones: 150 para el caso baseline, 30 para los demás.

Detección de puntos calientes en estructuras utilizando ondas guiadas

- Filtro: en nuestro caso, los experimentos se han realizado sin aplicar filtro

Señal utilizada

En los experimentos realizados en este proyecto se ha utilizado una señal senoidal modulada en amplitud por una ventana de Hamming de tal forma que la señal complete 5 oscilaciones (5 picos). Se ha decidido escoger esta señal porque ya ha sido utilizada en otros experimentos por el equipo CoDALab y ha presentado buenos resultados.

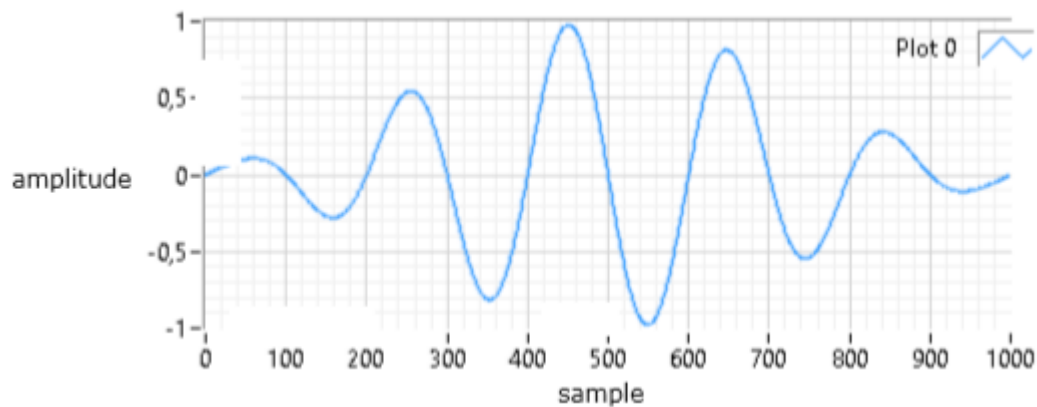


Figura 3.25: Señal utilizada en los experimentos.

Conexiones entre componentes

En la siguiente figura se muestra el diagrama de conexión entre el amplificador de señal utilizado, el generador de onda, el digitalizador de señales y el conmutador.

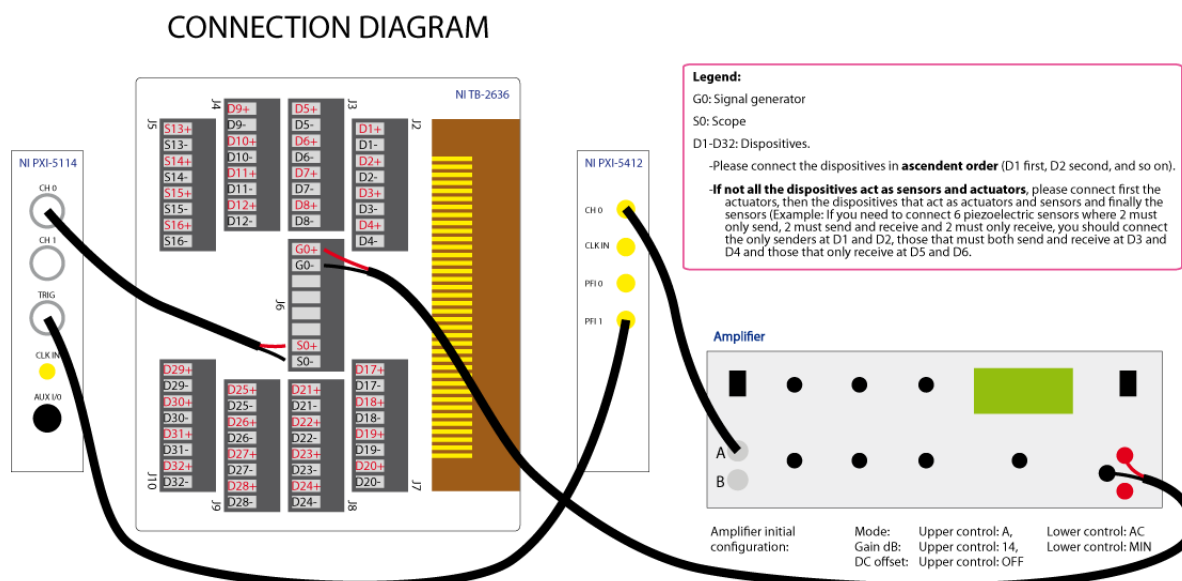


Figura 3.26: Diagrama de conexión entre componentes[Propiedad del equipo CoDALab].

3.4. Experimentos realizados

A lo largo del proyecto, se han realizado 9 casos de estudio diferentes. Para todos ellos, la disposición de los transductores piezoeléctricos se ha mantenido, teniendo así cinco PZT's: cuatro de ellos formando un cuadrado, los cuales actúan como sensor y el quinto en el centro del cuadrado, el cual tiene la función de actuador.

La diferencia entre cada caso es la localización del defecto (punto caliente), ya que este se ha generado en varios puntos de la placa para observar comportamientos diferentes en los sensores. Podemos dividir los casos en dos tipos:

1. El primer tipo pertenece a la realización de los experimentos sobre la estructura sana, es decir, sin haber calentado ningún punto (sin defecto). Para este caso se han realizado 150 experimentos. En base a los datos obtenidos en este experimento, se ha elaborado el modelo base de PCA. A este caso se le ha denominado "baseline".

Detección de puntos calientes en estructuras utilizando ondas guiadas

2. El segundo tipo se refiere a la realización de los experimentos cuando la placa presenta algún defecto. Existen 8 casos distintos, cada uno posee una localización del punto caliente diferente. Así mismo, a cada PZT se le ha asociado un número: al que sustenta la función de actuador se le ha asignado el número cero, los sensores se han numerado del uno al cuatro. Si unimos con una línea imaginaria cada piezoeléctrico con sus adyacentes se obtienen 8 “caminos” o “regiones” diferentes en la placa. A cada región se le ha nombrado según el número de los PZT que la forman y, según donde se genere el punto caliente nos encontraremos en una región u otra. Por ejemplo, si calentamos entre el actuador y el sensor 1, se está calentando en la región 0-1, por lo tanto nos encontramos en el caso “HS 0-1” (Hot spot 0-1). Para cada uno de estos ocho casos se han realizado 30 experimentos.

A continuación se encuentran las 8 imágenes captadas con la cámara termográfica para cada uno de los 8 casos con defecto. En la imagen para el caso HS 0-1 se puede apreciar aproximadamente la posición del actuador y de los sensores.

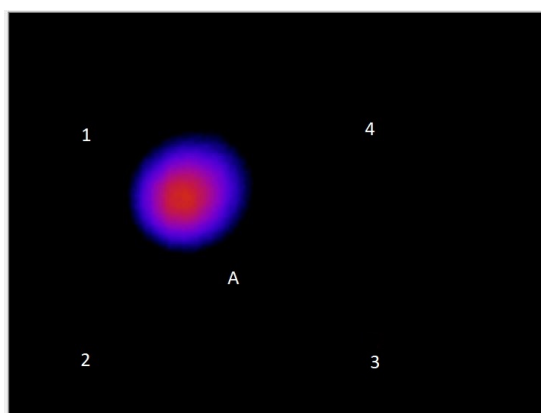


Figura 3.27: Hot Spot 0-1.

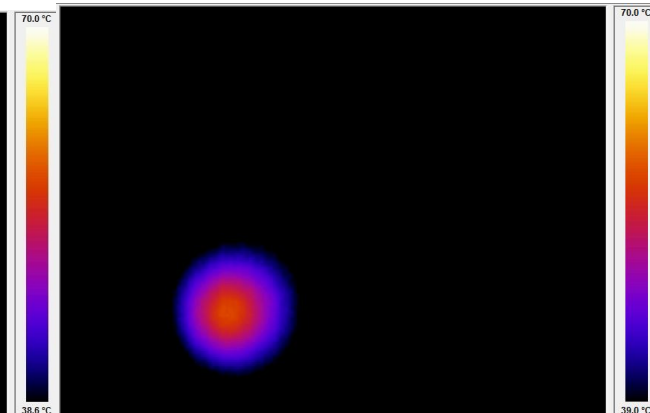


Figura 3.28: Hot Spot 0-2.

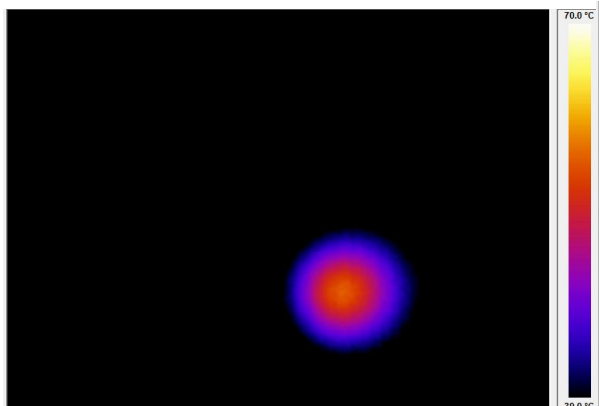


Figura 3.29: Hot Spot 0-3.

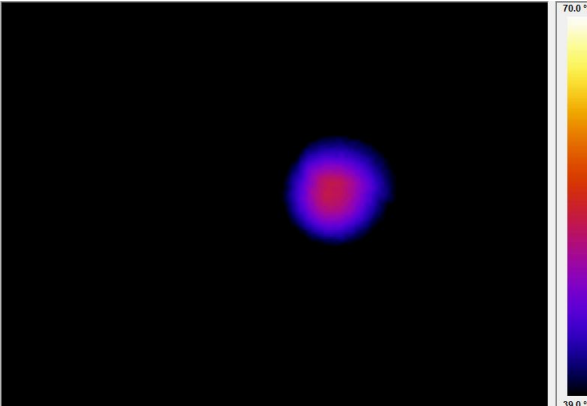


Figura 3.30: Hot Spot 0-4.

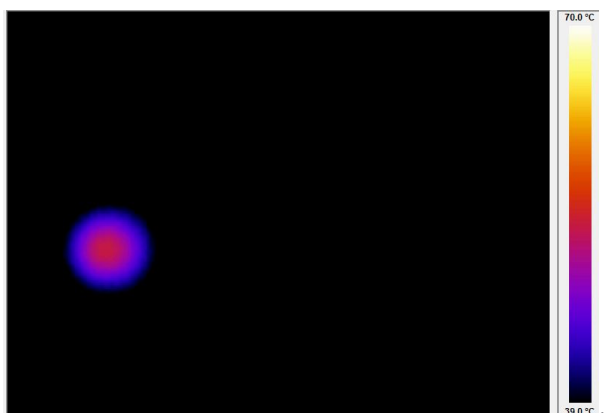


Figura 3.31: Hot Spot 1-2.

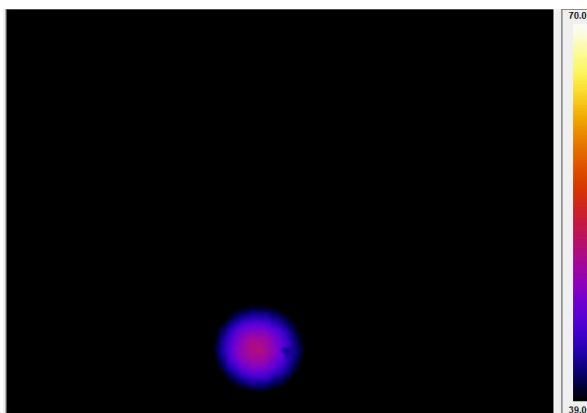


Figura 3.32: Hot Spot 2-3.

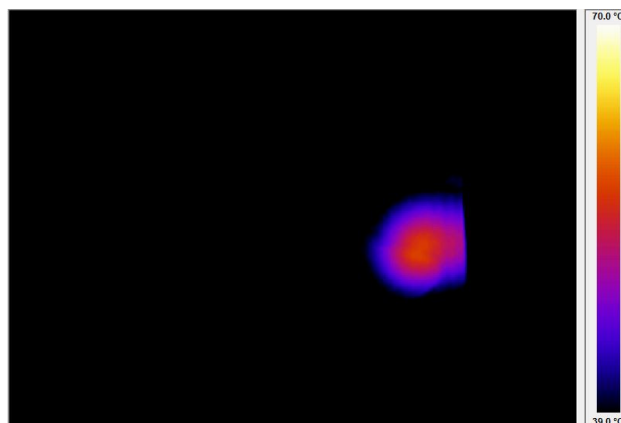


Figura 3.33: Hot Spot 3-4.

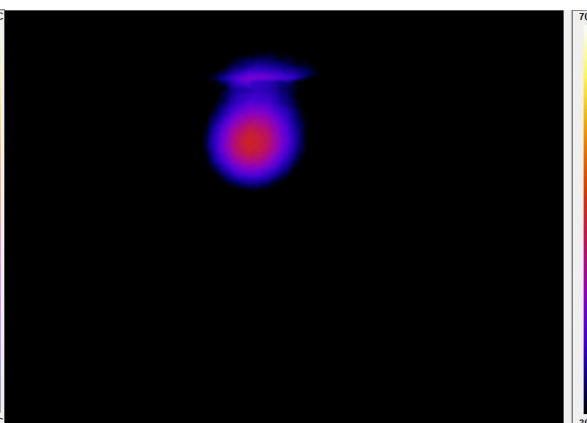


Figura 3.34: Hot Spot 4-1.

3.5. Organización de los datos en R

El fichero de datos sobre el que se trabaja en el presente proyecto se ha obtenido a partir de los archivos pertenecientes a cada experimento realizado. Estos archivos de datos se han juntado en uno solo creando un *data frame* para que sea más fácil su manipulación en R.

Un data frame al igual que una matriz, representa una estructura rectangular de datos, es decir, almacenan datos en filas y columnas. La diferencia entre ambos, es que las matrices almacenan un único tipo de datos mientras que los data frames pueden contener distintos tipos.

El data frame generado contiene 99000 observaciones (filas) y 154 variables (columnas).

Las filas están formadas por:

- 2750 muestras por cada señal (inicialmente se tenían 6875 muestras, pero al cortarlas las muestras se han reducido a 2750)
- 4 sensores
- Total de casos: 8 defectos + baseline

Multiplicando cada una de las variables anteriores se obtiene: $2750 * 4 * 9 = 99000$ filas.

Las columnas están formadas por:

- Defecte: Contiene los nombres de cada caso repetidas tantas veces como experimentos se hayan realizado Baseline, HS 0-1, HS 0-2...
- Actuator: número del actuador
- Sensor: número del sensor
- Time: tiempo de la señal. Este no empieza en 0, debido al corte de las señales.
- Experimentos O1-O150: 150 columnas, puesto que en el baseline se hicieron 150 experimentos mientras que para el resto solo 30. Las observaciones que correspondan al baseline tienen datos en las 150 columnas mientras que para el resto, las columnas del O31 al O150 no tienen datos (NA)

4. Análisis de componentes principales (ACP)

4.1. Conceptualización

Más conocido por sus siglas en inglés como Principal Component Analysis (PCA). Es un procedimiento estadístico que convierte un conjunto de variables posiblemente correlacionadas (cada una de las cuales toma varios valores) en un conjunto de datos no correlacionados linealmente llamados Componentes Principales.

También podemos entender este procedimiento como eliminar información redundante entre las dimensiones, reduciendo así la dimensionalidad de los datos.

4.1.2. Escalado o normalización de los datos

Por definición, los Componentes Principales son vectores que cruzan el origen. Si los datos no están centrados dimensionalmente (las columnas tienen una media de cero), el método no podrá obtener la mejor línea de regresión a lo largo de los datos. Por esta razón, las variables deben normalizarse.

Esto se recomienda especialmente cuando las variables se miden en diferentes escalas, ya que el método de PCA es sensible a estas diferencias. Otro caso en el que se deberían normalizar los datos es cuando la media y/o la desviación estándar de las variables son muy diferentes.

El objetivo es hacer que las variables sean comparables. En general, las variables se escalan para tener una media de cero y una varianza de la unidad.

Al escalar las variables, los datos pueden ser tratados como:

$$\mu_{v_j} = \frac{1}{n} \sum_{i=1}^n x_{ij} \quad (2)$$

$$\sigma_{v_j}^2 = \frac{1}{n-1} \sum_{i=1}^n (x_{ij} - \mu_{v_j})^2 \quad (3)$$

$$\underline{x}_{ij} = \frac{X_{ij} - \mu_{v_j}}{\sigma_{v_j}} \quad (4)$$

Donde μ_v y $\sigma_{v_j}^2$ son, respectivamente, la media y la varianza de las medidas del sensor j (v_j) y \underline{x}_{ij} es la muestra reescalada.

En lo que sigue de documento, las muestras reescaladas se mostrarán sin notación de barras por simplicidad.

4.1.2. Covarianza y correlación

Para la medición de la cantidad de información incorporada en una componente de una variable X , la cual contiene la medición tomada por m sensores en n tomas de datos, se utiliza su varianza. Es decir, cuanto mayor sea su varianza, mayor es la cantidad de información que lleva incorporada dicha componente. En caso de que su varianza sea 0, quiere decir que no hay información presente.

En cambio, si evaluamos más de un componente, el resultado del estudio es representado por una matriz de covarianzas.

$$\mathbf{C}_X = \frac{1}{n-1} \mathbf{X}^T \mathbf{X} = \frac{1}{n-1} \begin{pmatrix} \mathbf{v}_1^T \mathbf{v}_1 & \mathbf{v}_1^T \mathbf{v}_2 & \dots & \mathbf{v}_1^T \mathbf{v}_j & \dots & \mathbf{v}_1^T \mathbf{v}_m \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \mathbf{v}_j^T \mathbf{v}_1 & \mathbf{v}_j^T \mathbf{v}_2 & \dots & \mathbf{v}_j^T \mathbf{v}_j & \dots & \mathbf{v}_j^T \mathbf{v}_m \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \mathbf{v}_m^T \mathbf{v}_1 & \mathbf{v}_m^T \mathbf{v}_2 & \dots & \mathbf{v}_m^T \mathbf{v}_j & \dots & \mathbf{v}_m^T \mathbf{v}_m \end{pmatrix} \quad (5)$$

La matriz C_x es cuadrada, simétrica y de tamaño $m \times m$. Esta matriz mide, para cada par de sensores, el grado de relación lineal existente entre los datos tomados por los mismos.

Los elementos de la diagonal de la matriz C_x son las varianzas de los datos tomados por cada uno de los sensores.

$$\sigma_{v_j}^2 = \frac{1}{n-1} v_j^T v_j = \frac{1}{n-1} \sum_{i=1}^n x_{ij}^2 \quad (6)$$

Los elementos que no están en la diagonal, son las covarianzas entre los datos tomados por cada par de sensores diferentes.

$$\sigma_{v_j, v_k}^2 = \frac{1}{n-1} v_j^T v_k = \frac{1}{n-1} \sum_{i=1}^n x_{ij} x_{ik} \quad (7)$$

Podemos observar que un valor alto de covarianza entre un determinado par de sensores implica la existencia de una alta redundancia entre los datos contenidos en dicho par de sensores.

Como la matriz C_x es cuadrada y simétrica, es diagonalizable. Si pudiésemos expresar los datos contenidos en X de forma tal que la matriz C_x fuera diagonal, estaríamos expresando los datos sin redundancia alguna entre pares de variables, condensando así el total de información contenida en la matriz X .

4.1.3. Transformación de la matriz de datos

Consideremos una matriz P de transformación lineal y de tamaño $m \times n$ utilizada para transformar la matriz de datos X de la forma siguiente:

$$T = XP \quad (8)$$

Siendo la matriz de covarianzas de T diagonal, es decir, que la covarianza entre cualquier par de columnas de la misma sea nula:

Detección de puntos calientes en estructuras utilizando ondas guiadas

$$C_T = \frac{1}{n-1} T^T T = \text{diagonal} \quad (9)$$

Si combinamos las dos ecuaciones comentadas anteriormente (8) y (9):

$$C_T = \frac{1}{n-1} P^T X^T X P = P^T C_X P \quad (10)$$

Como se ha comentado en el apartado anterior, C_X es simétrica. Tiene m valores propios reales λ_j y m vectores propios ortonormales p_j , que forman una base de un espacio m dimensional.

Utilizando como matriz de transformación una matriz que tiene por columnas los vectores propios de C_X :

$$P = (p_1 | p_2 | \dots | p_j | \dots | p_m) \quad (11)$$

Podemos satisfacer la siguiente propiedad:

$$C_X P = P \Lambda \quad (12)$$

Donde $\Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_m)$. Sustituyendo (12) en (10) se puede apreciar que se satisface a la condición (12):

$$C_T = P^T P \Lambda = \Lambda \quad (13)$$

Si escribimos la transformación $T = X P$ con más detalle:

$$(t_1 | t_2 | \dots | t_j | \dots | t_m) = \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1j} & \dots & x_{1m} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ x_{i1} & x_{i2} & \dots & x_{ij} & \dots & x_{im} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ x_{n1} & x_{n2} & \dots & x_{nj} & \dots & x_{nm} \end{pmatrix} \times (p_1 | p_2 | \dots | p_j | \dots | p_m) \quad (14)$$

Cada columna en la matriz T puede ser expresada como vector:

$$t_j = X p_j \quad (15)$$

Las varianzas de estos vectores son:

$$\sigma_{t_j}^2 = \frac{1}{n-1} t_j^T t_j = \frac{1}{n-1} (X p_j)^T (X p_j) = p_j^T C_x p_j = \lambda_j \quad (16)$$

Mientras que sus covarianzas son nulas:

$$\sigma_{t_j, t_k}^2 = \frac{1}{n-1} t_j^T t_k = \frac{1}{n-1} (X p_j)^T (X p_k) = p_j^T C_x p_k = \lambda_j p_j^T p_k = 0 \quad (17)$$

Así mismo, no existe correlación entre dos vectores columna cualesquiera de la matriz transformada T y sus variaciones están dadas por los valores propios de la matriz de covarianzas C_x de los datos originales. Debido a que un mayor valor de varianza implica que el contenido de información es mayor, los vectores propios que forman la matriz de transformación P se clasifican según sus valores propios en orden descendente, y se denominan Componentes Principales del conjunto de datos. De esta forma, los vectores columna de la matriz transformada T quedan ordenados según la cantidad de información que contienen y se denominan *scores*.

4.1.4.Reducción de la dimensión de la matriz de datos transformada

Debido a que los vectores propios se ordenan de acuerdo a su cantidad de información, es posible reducir la dimensionalidad del conjunto de datos al escoger sólo un número reducido r de componentes principales, que corresponden a los primeros valores propios. De esta forma podemos obtener la siguiente matriz de transformación $m \times r$:

$$P = (p_1 | p_2 | \dots | p_r) \quad (18)$$

De tal forma que la matriz transformada T contenga gran parte de la información originalmente contenida en la matriz X , ofreciendo una versión dimensionalmente reducida de la misma.

En caso de disponer de todos los componentes principales, es posible invertir la proyección y recuperar los datos originales según la ecuación:

$$X = T P^T \quad (19)$$

En caso de no disponer de todos los componentes principales, es posible recuperar la matriz X pero con cierta pérdida de información a través de la transformación. Esta versión de la matriz X es llamada \hat{X} :

$$\hat{X} = T P^T = X (P P^T) \quad (20)$$

Mediante simples manipulaciones, podemos llegar a la siguiente expresión de la matriz X :

$$X = \hat{X} + \tilde{X} \quad (21)$$

$$\hat{X} = X (P P^T) \quad (22)$$

$$\tilde{X} = X (I - P P^T) \quad (23)$$

Donde \hat{X} es la proyección de la matriz de datos X sobre las componentes principales seleccionadas y \tilde{X} es la proyección sobre los componentes residuales.

4.1.5. Métodos de cálculo de la matriz de transformación lineal P

Existen varios métodos para calcular la matriz de transformación lineal P y la matriz transformada T correspondiente. A continuación explicaremos dos de los métodos más utilizados.

Uno de los métodos es por medio de la descomposición de valores singulares (Singular Value Decomposition, SVD) de la matriz de covarianzas (ecuación (5)). Los subespacios en PCA se definen por los vectores propios y los valores propios de la matriz de covarianzas de tal forma que se satisface la ecuación (12). Así mismo, el método SVD nos permite calcular la matriz de transformación P de modo que tenga como columnas los vectores propios de la matriz de covarianzas Cx . Dichas columnas están ordenadas según los valores propios en orden descendente.

El otro método es utilizar el algoritmo NIPALS (Nonlinear Iterative Partial Least Square), el cual es más eficiente y preciso para matrices grandes, pero más lento que SVD. Considerando cada vector columna t como las columnas de la matriz transformada T y, cada vector columna p como las columnas de la matriz de transformación P , el objetivo del algoritmo NIPALS es encontrar por iteraciones los vectores t y p de acuerdo con la ecuación (8). Esta relación puede ser expresada de forma vectorial como $t = X * p$.

El algoritmo NIPALS se puede describir como [13]:

1. Inicializar el vector t utilizando la variable (columnas en X) con la mayor varianza (t_{old}).
2. Estimar el vector de los loadings correspondiente (columna en P) en el vector de los scores t_{old} utilizando la expresión $p = (X^T t_{old}) / (t_{old}^T t_{old})$.
3. Normalizar el vector de los loadings a un vector unitario de la manera: $p = p / \|p\|$.

Detección de puntos calientes en estructuras utilizando ondas guiadas

4. Proyectar X en el vector p para encontrar una nueva estimación del vector t (t_{new}):

$$t_{new} = (Xp) / (p^T p).$$

5. Si $d = \frac{(t_{new} - t_{old})'(t_{new} - t_{old})}{(mt'_{new} - t_{old})} > 10^{-10}$, volvemos al paso número 2. Sino t_{new} es un score. En la

ecuación de convergencia d , m es el número de sensores.

6. Se resta la varianza explicada para el vector de loadings actual de X , $X_{new} = X - t p^T$, restando los datos proyectados de X .
7. Se vuelve al paso número 1 con tal de estimar otros componentes.

4.1.6. Ejemplo aplicación PCA

Seguidamente se muestra un ejemplo de aplicación del análisis de componentes principales en la simplificación de un conjunto de datos.

Supongamos que disponemos de las calificaciones de 12 alumnos del último examen (evaluado sobre 100 puntos) que han realizado y del tiempo que ha dedicado cada alumno a la preparación del examen.

Alumno	Tiempo de estudio (h)	Calificación
1	9	39
2	15	56
3	25	93
4	14	61
5	10	50
6	18	75
7	0	32
8	16	85
9	5	42
10	19	70
11	16	66
12	20	80

El objetivo es reducir la dimensión de los datos (dos, en este caso) a una sola dimensión o componente. Ordenando los datos de forma matricial obtenemos:

$$X = \begin{pmatrix} 9 & 39 \\ 15 & 56 \\ 25 & 93 \\ 14 & 61 \\ 10 & 50 \\ 18 & 75 \\ 0 & 32 \\ 16 & 85 \\ 5 & 42 \\ 19 & 70 \\ 16 & 66 \\ 20 & 80 \end{pmatrix}$$

Si tomamos como dimensiones las calificaciones y el tiempo de estudio de cada alumno, obtenemos un espacio bidimensional. Podemos representar los datos anteriores en un gráfico situando las horas de estudio en el eje de abscisas y la calificación de cada alumno en el eje de ordenadas

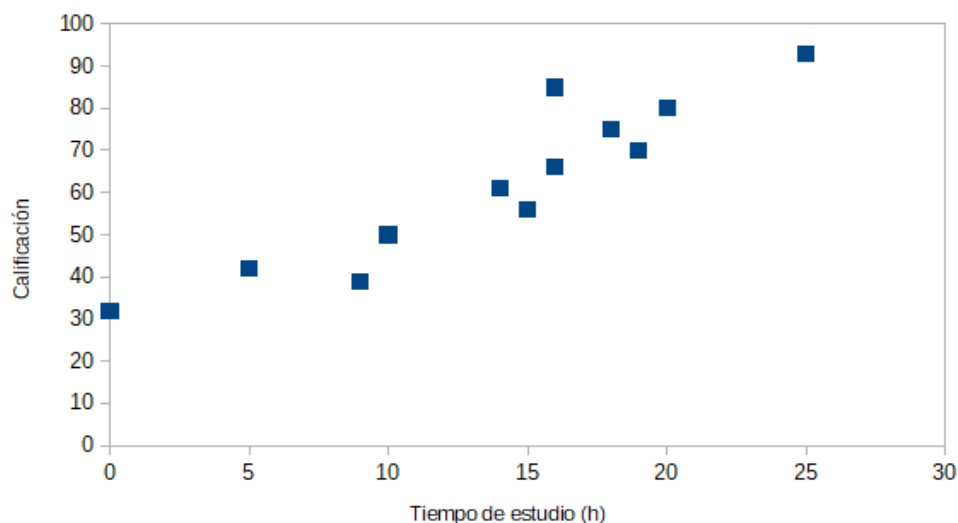


Figura 4.1: Representación de los datos originales para el ejemplo de PCA.

Podemos observar la existencia de correlación entre los datos, lo que es algo lógico ya que normalmente la nota obtenida en un examen depende del tiempo que le hayas dedicado a prepararlo.

Así mismo podríamos decir que en cuanto mayor es el número de horas de estudio, mayor es la nota obtenida.

Paso 1: Normalización

Como se ha comentado anteriormente, los datos se deben normalizar de forma que cada dimensión tenga media 0 y varianza 1. Haciendo uso de las ecuaciones (2), (3) y (4), obtenemos los siguientes valores (los datos reescalados se muestran sin notación de barras para mayor simplicidad):

$$X = \begin{pmatrix} -1,0226 & 0,0218 \\ -0,8123 & 0,4672 \\ -0,4616 & 1,4367 \\ -0,8473 & 0,5983 \\ -0,9876 & 0,3100 \\ -0,7071 & 0,9651 \\ -1,3382 & -0,1616 \\ -0,7772 & 1,2271 \\ -1,1629 & 0,1004 \\ -0,6720 & 0,8341 \\ -0,7772 & 0,7293 \\ -0,6370 & 1,0961 \end{pmatrix} = (v_1 | v_2)$$

A continuación podemos ver los datos reescalados de forma gráfica.

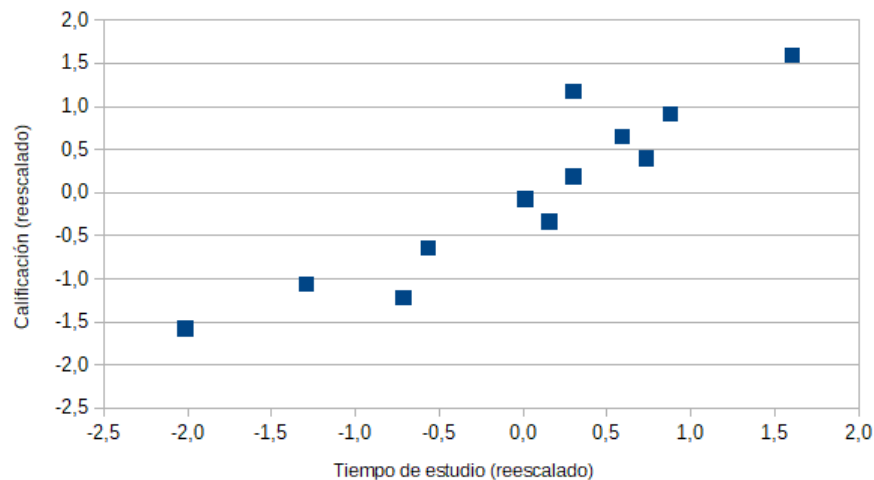


Figura 4.2: Datos del ejemplo PCA normalizados.

Paso 2: Matriz de covarianzas

Recordando la ecuación (5), calculamos la matriz de covarianzas como $C_x = \frac{1}{n-1} X^T X$ y obtenemos:

$$C_x = \begin{pmatrix} 1 & 0,9252 \\ 0,9252 & 1 \end{pmatrix}$$

La matriz resultante es cuadrada y simétrica. La diagonal representa la varianza de cada dimensión. Podemos observar que la covarianza es positiva, por lo que las dos dimensiones aumentan juntas.

Podemos verificar los resultados obtenidos calculando directamente la covarianza entre v_1 y v_2 a través de la ecuación (7):

$$\sigma_{v_j, v_k}^2 = \frac{1}{12-1} v_1^T v_2 = 0,9252$$

Paso 3: Cálculo de valores propios y vectores propios de la matriz C_x

Los vectores propios de la matriz C_x son:

$$v_{ep_1} = \begin{pmatrix} 0,7071 \\ 0,7071 \end{pmatrix}$$

$$v_{ep_2} = \begin{pmatrix} -0,7071 \\ 0,7071 \end{pmatrix}$$

Y los valores propios asociados a estos vectores son:

$$\lambda_1 = 1,9252$$

$$\lambda_2 = 0,0748$$

Detección de puntos calientes en estructuras utilizando ondas guiadas

Una vez encontrados los valores y vectores propios podemos construir la matriz de transformación lineal P. Las columnas de la matriz P están formadas componentes principales de la matriz C_x , ordenados según el valor propio asociado a cada vector en orden descendente.

$$P = \begin{pmatrix} 0,7071 & -0,7071 \\ 0,7071 & 0,7071 \end{pmatrix}$$

Paso 4: Construcción de la matriz T

A continuación calculamos la matriz T, resultante de la transformación lineal $T = XP$

$$T = XP = \begin{pmatrix} -0,7077 & 0,7385 \\ -0,2440 & 0,9047 \\ 0,6895 & 1,3423 \\ -0,1761 & 1,0222 \\ -0,4791 & 0,9175 \\ 0,1824 & 1,1824 \\ -1,0605 & 0,8320 \\ 0,3181 & 1,4172 \\ -0,7513 & 0,8933 \\ 0,1146 & 1,0650 \\ -0,0339 & 1,0652 \\ 0,3246 & 1,2255 \end{pmatrix}$$

La matriz T contiene los datos de origen expresados en una base ortonormal diferente a la original (la base formada por los componentes principales de la matriz \underline{X}). La siguiente gráfica contiene los datos de la matriz T, donde el componente principal 1 se encuentra en el eje de abscisas y el componente principal 2 en el eje de ordenadas:

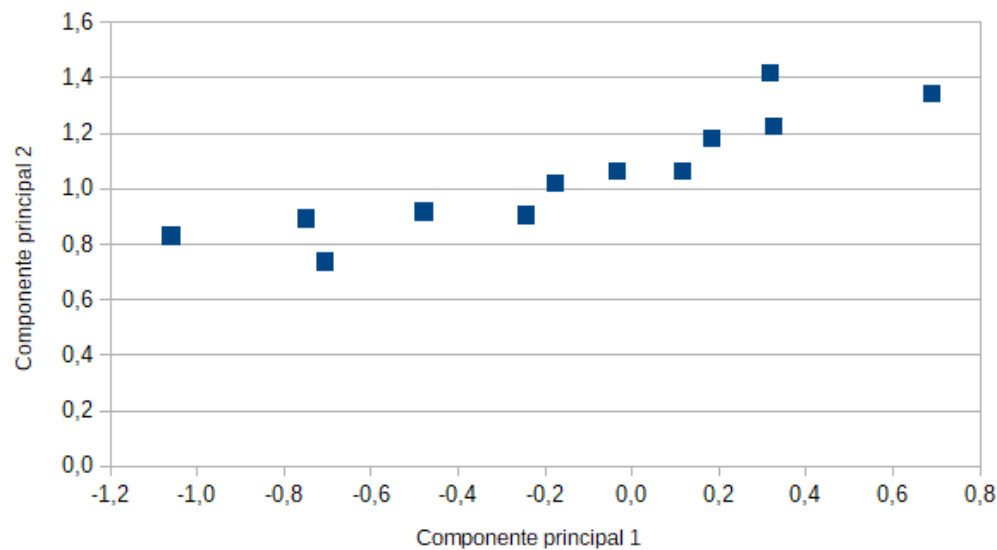


Figura 4.3: Primer y segundo componente principal para el ejemplo de PCA.

Se puede observar que la proyección de los datos sobre el primer componente principal contiene más información (mayor varianza) que la proyección sobre segundo componente principal. Esto ya lo habíamos visto antes, ya que nos lo señalan los valores propios.

Si deseamos obtener una representación dimensionalmente simplificada de los datos, descartamos los valores correspondientes al segundo componente principal, ya que es el que contiene menos información. Así mismo, obtenemos una versión reducida de la matriz T que contiene los datos de la primera componente principal:

$$T_1 = \begin{pmatrix} -0,7077 \\ -0,2440 \\ 0,6895 \\ -0,1761 \\ -0,4791 \\ 0,1824 \\ -1,0605 \\ 0,3181 \\ -0,7513 \\ 0,1146 \\ -0,0339 \\ 0,3246 \end{pmatrix}$$

Paso 5: Recuperación de la información original

En base a los datos contenidos en la matriz T reducida, ahora se desea recuperar los datos originales normalizados. La mayoría de los datos se recuperarán sin grandes errores visibles, manteniendo así la mayor cantidad de información. Para ello se calculará la matriz \hat{X} a partir de la ecuación (20).

4.2. Índices de detección de anomalías

El método de análisis PCA puede usarse para detectar comportamiento anormal en procesos o sistemas. Dos estadísticos conocidos comúnmente utilizados para este objetivo son: Q-statistic y T^2 -statistic. El primer método se basa en analizar los datos residuales (matriz \tilde{X}) para representar la variabilidad de los datos proyectados en el espacio residual. El segundo se basa en analizar la matriz de puntos T para verificar la variabilidad de los datos proyectados en el nuevo espacio de componentes principales. Estos dos métodos se basan en la suposición (generalmente derivada del teorema central del límite) que el proceso sigue aproximadamente una distribución normal multivariada.

Distancia de Mahalanobis

Antes de definir los dos métodos estadísticos comentados anteriormente, es necesario comentar la distancia de Mahalanobis. Esta es una medida de distancia entre un punto P y una distribución D.

Supongamos que queremos medir la distancia entre dos puntos $x_1(x_{11}, x_{12})$ y $x_2(x_{21}, x_{22})$. La distancia Euclídea sería:

$$d(x_1, x_2) = \sqrt{(x_{11} - x_{21})^2 + (x_{12} - x_{22})^2} \quad (24)$$

Puesto que las dos dimensiones no deben tener la misma escala necesariamente, son normalizadas utilizando la desviación estándar

$$d(x_1, x_2) = \sqrt{\left(\frac{x_{11} - x_{21}}{\sigma_1}\right)^2 + \left(\frac{x_{12} - x_{22}}{\sigma_2}\right)^2} \quad (25)$$

Donde sigma (σ) es la desviación estándar de cada dimensión. Representando en forma vectorial:

$$d(x_1, x_2) = \sqrt{(x_1 - x_2)^T S^{-1} (x_1 - x_2)} \quad (26)$$

Donde S es la matriz diagonal con la desviación estándar en los elementos diagonales. Dado que las dimensiones puede no estar incorrelacionadas, esta distancia puede ser corregida utilizando la matriz de covarianzas C_x .

$$d(x_1, x_2) = \sqrt{(x_1 - x_2)^T C_x^{-1} (x_1 - x_2)} \quad (27)$$

Obteniendo así la expresión de la distancia de Mahalanobis. Podemos observar que lo que diferencia esta distancia de la Euclídea, es que Mahalanobis tiene en cuenta la correlación entre las variables.

Estadístico T^2

Basado en la distancia de Mahalanobis, donde una de las variables se compara con la media de la distribución base.

$$T^2 = (N - 1) (\bar{X} - \mu_0)^T S^{-1} (\bar{X} - \mu_0) \quad (28)$$

El estadístico T^2 es una generalización de *Student's t-statistic* que es usada en pruebas de hipótesis multivariable. Describe una medida de la variación en cada muestra dentro del modelo (indica que tan lejos del modelo se encuentra).

El estadístico T^2 de Hotellings, se define como la suma de los datos normalizados al cuadrado. Consideramos x_i como el vector fila m que representa las medidas de todos los sensores en el experimento i . Llamamos t_i al vector fila r (fila i en la matriz T), que es la proyección del experimento x_i en el nuevo espacio. Ambos se relacionan como $t_i = x_i P$.

El estadístico T^2 de la muestra (o experimento) i se define como:

$$T_i^2 = \sum_{j=1}^r \frac{t_{ij}^2}{\lambda_j} = t_i \Lambda^{-1} t_i^T = x_i (P \Lambda^{-1} P^T) x_i^T \quad (29)$$

Estadístico Q

El estadístico Q expresa el cambio de los eventos que no están explicados por el modelo de componentes principales. Es una medida de la diferencia entre una muestra y su proyección en el modelo. Considerando una muestra o experimento i del vector x_i se define como:

$$Q_i \|\tilde{x}\|^2 = \tilde{x}_i \tilde{x}_i^T = x_i (I - P P^T) x_i^T \quad (30)$$

Donde x_i es la proyección en el subespacio residual.

Normalmente Q-statistic es más sensible que T^2 -statistic. Esto se debe a que Q es muy pequeño y entonces cualquier cambio, por mínimo que sea, es observable. T^2 tiene una gran varianza por lo que requiere un gran cambio en las características del sistema para ser detectable [14].

4.3. PCA aplicado al SHM

En los apartados “Análisis de componentes principales” e “índices de detección de anomalías” comentados anteriormente, la base de PCA se ha considerado como una matriz X que contiene mediciones de varias variables para diferentes ensayos experimentales. Asumiendo así que cada medición es una muestra individual en el conjunto de datos (un solo valor). En las aplicaciones de SHM (Structural Health Monitoring) basadas en señales captadas por transductores piezoeléctricos, estas mediciones no representan un solo valor, sino que son conjuntos de valores captados por cada uno de los sensores (una señal percibida por uno de los sensores, se compone de un conjunto de muestras que han sucedido en el tiempo).

Cuando se realiza la toma de datos sobre una estructura, se denomina fase de excitación a la toma de datos llevada a cabo durante la excitación de la estructura por parte de un determinado actuador, y a la captación de la respuesta dinámica por parte del resto de transductores. Suponiendo que disponemos de una configuración con 4 transductores piezoeléctricos, la fase de excitación 1 corresponde a la emisión de la señal por parte del transductor 1 (actuador) y a la recepción de la señal por parte de los transductores 2,3 y 4 (sensores). Con el fin de obtener varios conjuntos de datos sobre los que efectuar análisis estadísticos, cada fase de excitación es repetida varias veces.

Una vez realizada la toma de datos, el conjunto de estos se organiza en una matriz tridimensional X_{3D} (I experimentos $\times K$ muestras por experimento $\times J$ sensores). Cada rebanada frontal es una matriz bidimensional X que contiene una fila por cada repetición del experimento realizada y

Detección de puntos calientes en estructuras utilizando ondas guiadas

una columna por cada muestra de las que se compone una señal digitalizada, es decir, representa todas las muestras de un sensor.

Para poder considerar las correlaciones en el tiempo para cada señal y las correlaciones entre los sensores, y así poder aplicar PCA, la matriz tridimensional inicial se debe desplegar en una matriz bidimensional como se muestra en la siguiente figura.

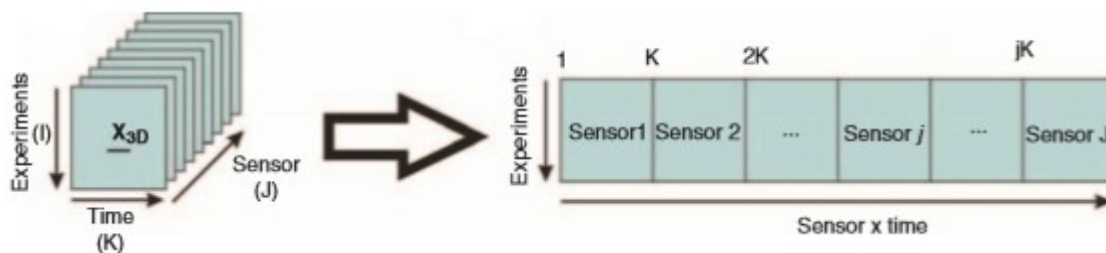


Figura 4.4: Despliegue de la matriz de datos inicial [15].

Escalado de datos

Como ya se ha comentado anteriormente, es necesario escalar los datos antes de aplicar PCA. Para este tipo de conjuntos de datos (matrices desplegadas), se presentan varios estudios diferentes de escalado de datos: escalamiento continuo (CS), escalado de grupos (GS) y autoescalado (AS). En el presente proyecto se han utilizado tanto escalado de grupos como autoescalado.

GS realiza un escalado basado en desviaciones estándares, se utiliza sobretudo cuando los datos constan de varios bloques de variables de igual tamaño. Cada muestra es escalada usando la media de todas las muestras tomadas por un mismo sensor en un mismo instante relativo de tiempo y la desviación estándar de todas las muestras tomadas por un mismo sensor.

$$\mu_{jk} = \frac{1}{I} \sum_{i=1}^I x_{ijk} \quad (31)$$

$$\mu_j = \frac{1}{IK} \sum_{i=1}^I \sum_{k=1}^K x_{ijk} \quad (32)$$

$$\sigma_j^2 = \frac{1}{IK} \sum_{i=1}^I \sum_{k=1}^K (x_{ijk} - \mu_j)^2 \quad (33)$$

$$z_{ij} = \frac{x_{ij} - \mu_{jk}}{\sigma_j} \quad (34)$$

AS en cambio, no utiliza la desviación estándar de todas las muestras tomadas por un sensor, sino que utiliza la desviación estándar de cada columna.

En las siguientes figuras se pueden apreciar las señales de 4 sensores, para el mismo experimento. En la primera figura aún no se le ha aplicado ningún escalado a los datos, mientras que en la segunda sí.

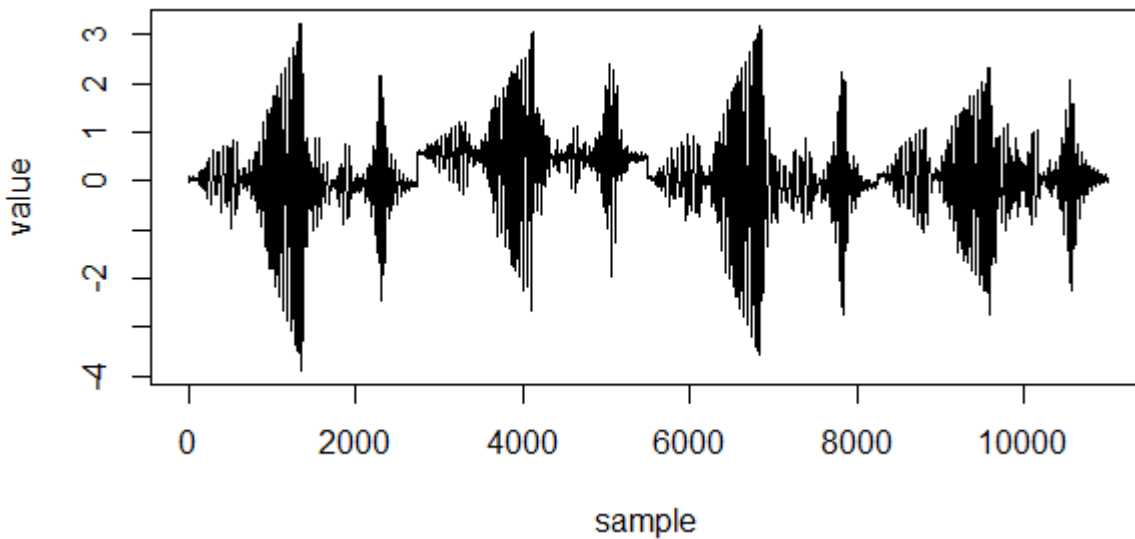


Figura 4.5: Señales de los datos sin escalar de 4 sensores para el mismo experimento.

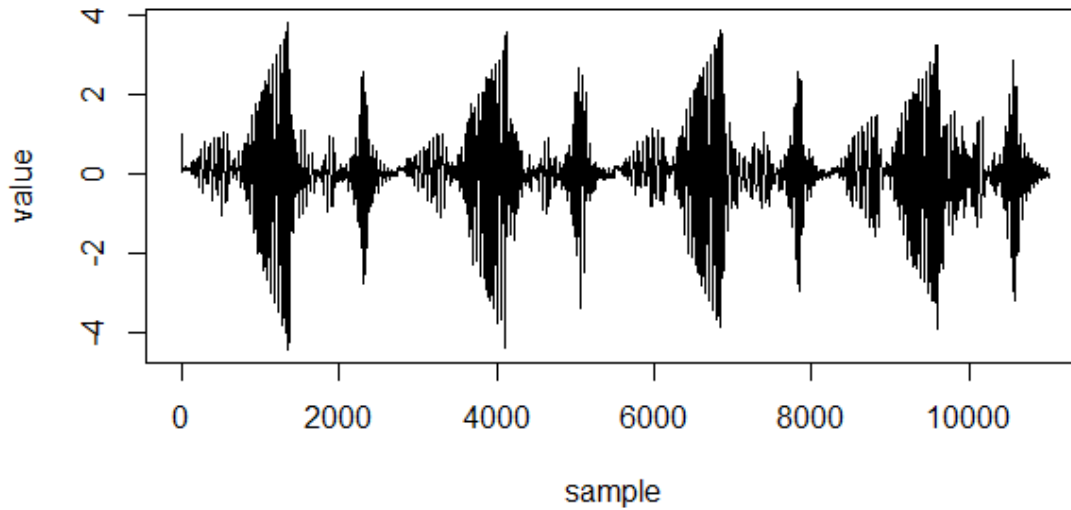


Figura 4.6: Señales de los datos escalados de 4 sensores para el mismo experimento.

Modelo base y diagnóstico

Para poder realizar un diagnóstico sobre una estructura, se debe construir una línea base a partir de datos del sistema funcionando en las condiciones normales deseadas. En nuestro caso, las condiciones deseadas son cuando la estructura es sana.

Así mismo, se debe aplicar PCA a la matriz que contiene todas las señales recopiladas de diferentes experimentos que utilizan la muestra base (estructura sana). Estas señales se organizan, tal y como se ha explicado en el apartado 4.3 del presente proyecto, obteniendo una matriz X . Esta matriz tiene una dimensión $I \times KJ$ (I experimentos, K muestras por experimento, J sensores). Una vez obtenida la matriz X , se aplica un escalado, siguiendo las ecuaciones anteriormente comentadas, para finalmente aplicar PCA sobre la matriz X escalada.

Construir una línea base, significa calcular la matriz de proyección P . Esta matriz será el modelo de la estructura no dañada que se utilizará en el diagnóstico de salud.

En lo que se refiere al diagnóstico, la cantidad de experimentos sobre esta estructura puede ser del número que se desee, pero el número de sensores y de muestras debe ser el mismo que se ha utilizado en el modelo base.

Una vez realizados los experimentos para el diagnóstico, se obtiene una nueva matriz X . Los datos de esta nueva matriz se normalizan con los mismos coeficientes que se utilizaron en los datos base y se proyectan sobre el modelo base, obteniéndose así los scores de dichos datos. Seguidamente se calculan los índices de daño (estadístico T^2 y estadístico Q) y se comparan con los valores del modelo base.

5. Análisis preliminar de datos

Antes de aplicar PCA, se ha realizado un análisis preliminar de los datos comparando las señales obtenidas en cada sensor para cada experimento.

Inicialmente se han comparado las señales de cada experimento para un solo un caso de estudio para todos los sensores. A continuación podemos observar un gráfico con las señales de los 150 experimentos del caso baseline para el sensor número 1:

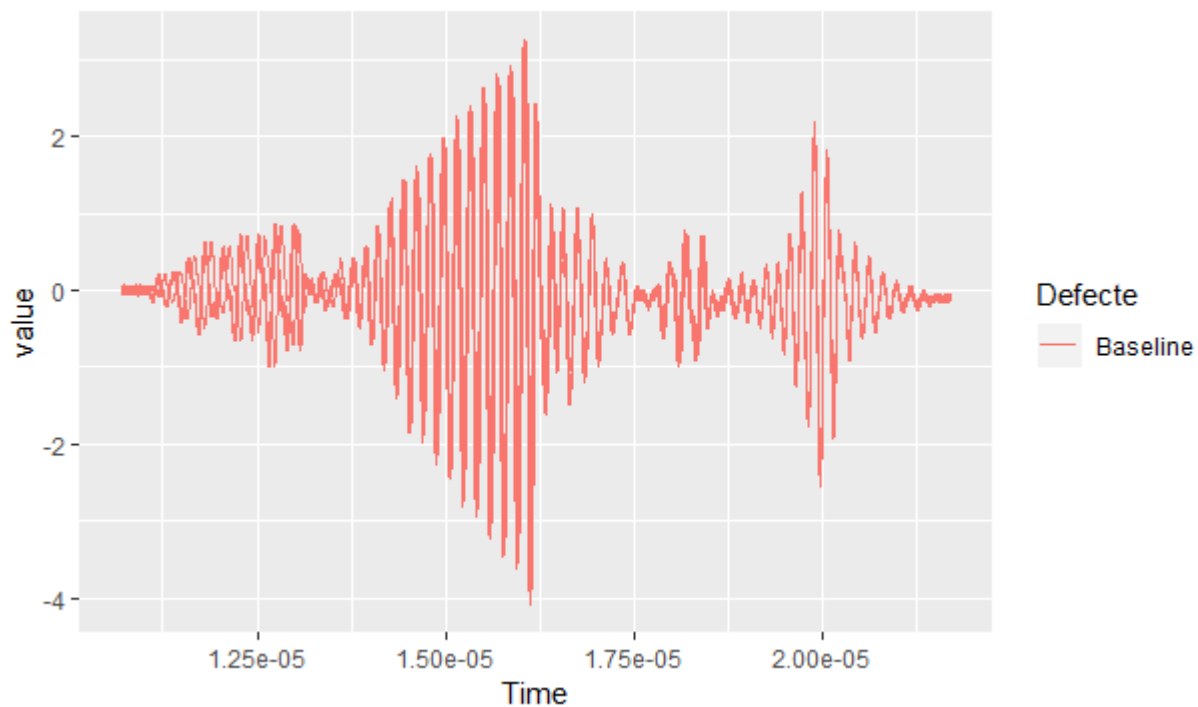


Figura 5.1: Señales de los experimentos del caso baseline sobre el mismo eje de tiempo.

En este gráfico todas las señales están graficadas sobre el mismo eje de tiempo para así poder ver si hay diferencias entre ellas. A simple vista se puede apreciar como las señales son semejantes en lo que se refiere a la forma y a su amplitud, pero no son idénticas.

Seguidamente se han comparado, para cada sensor, las señales obtenidas de los experimentos del caso baseline con las de los experimentos de los casos con defecto en la estructura. En el siguiente gráfico se puede ver una comparación para el sensor uno entre el caso baseline y el caso HS 0-2:

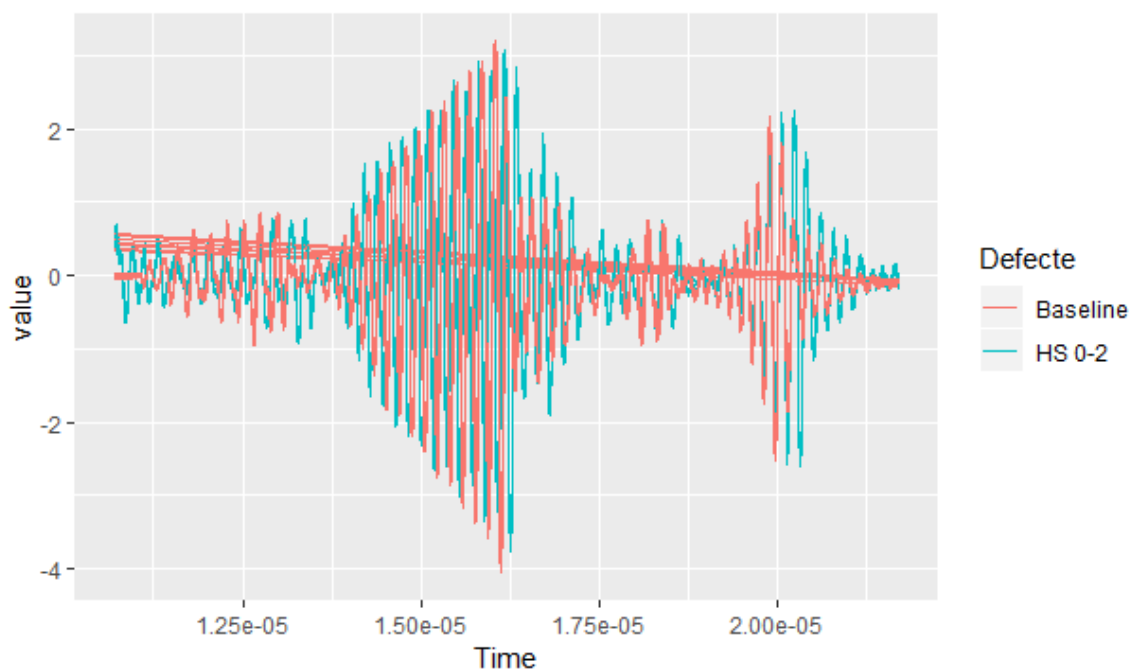


Figura 5.2: Comparación señales caso baseline y caso HS 0-2 sobre el mismo eje de tiempo.

Tan solo observando el gráfico, se puede apreciar ciertas diferencias entre ambos casos. Las señales para el caso “HS 0-2” presentan una amplitud mayor aunque, la forma de las señales sigue siendo muy parecida a las del baseline.

A continuación se pueden observar una serie de gráficos, estos nos muestran todos los experimentos graficados para cada uno de los casos y separados para cada sensor.

Detección de puntos calientes en estructuras utilizando ondas guiadas

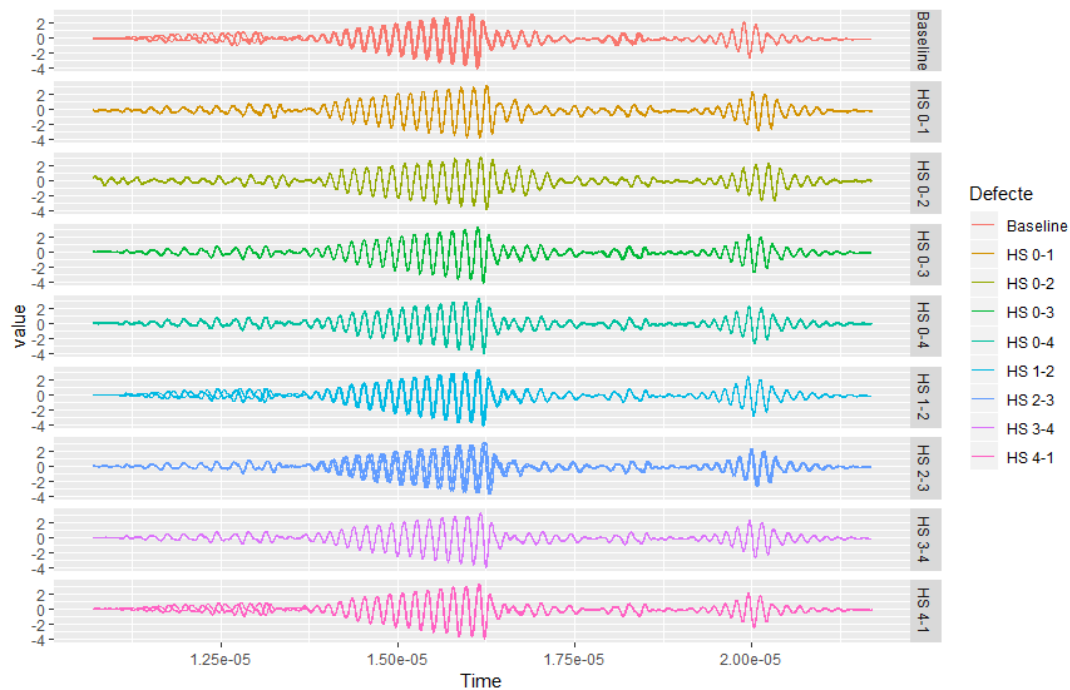


Figura 5.3: Señales de los experimentos de todos los casos para el sensor 1.

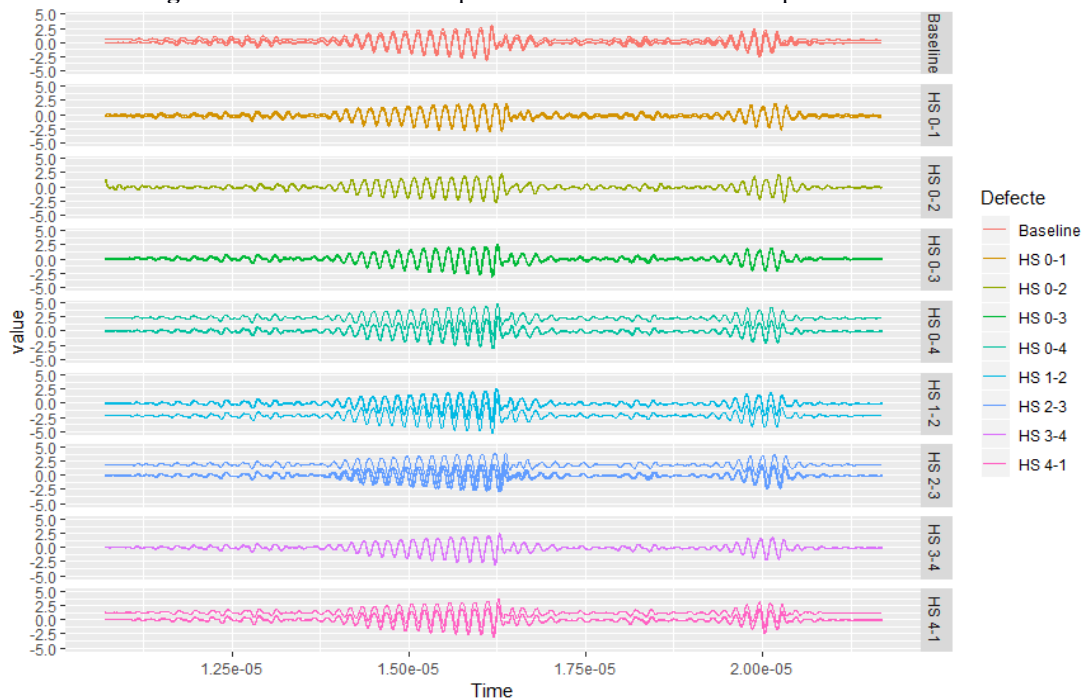


Figura 5.4: Señales de los experimentos de todos los casos para el sensor 2.

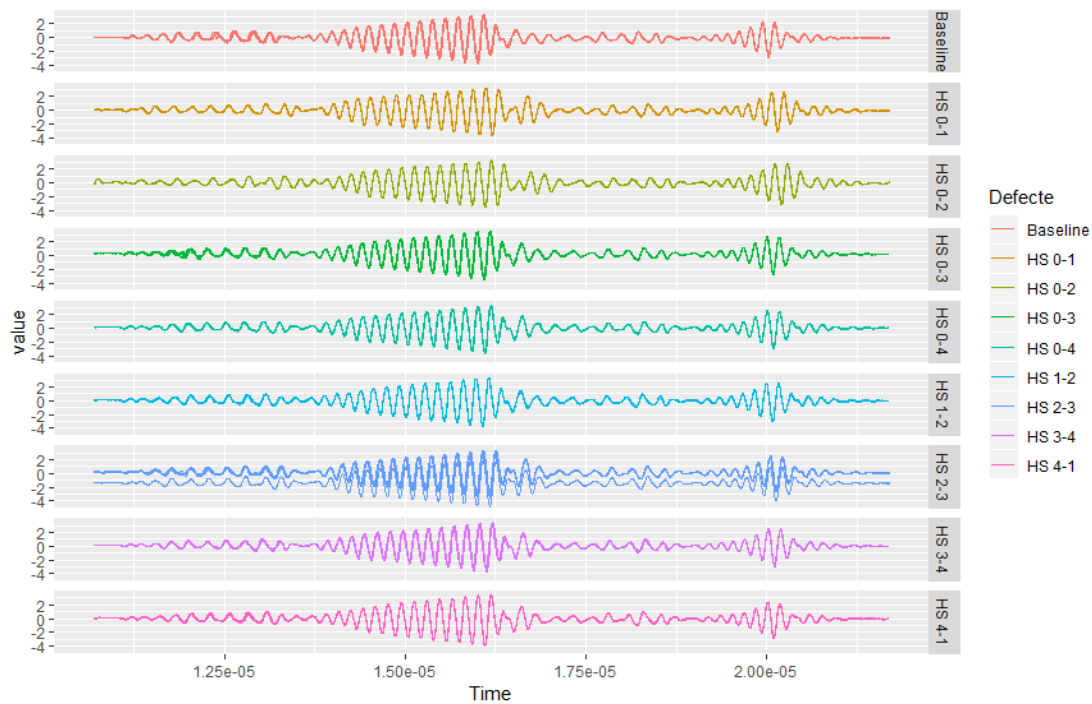


Figura 5.5: Señales de los experimentos de todos los casos para el sensor 3.

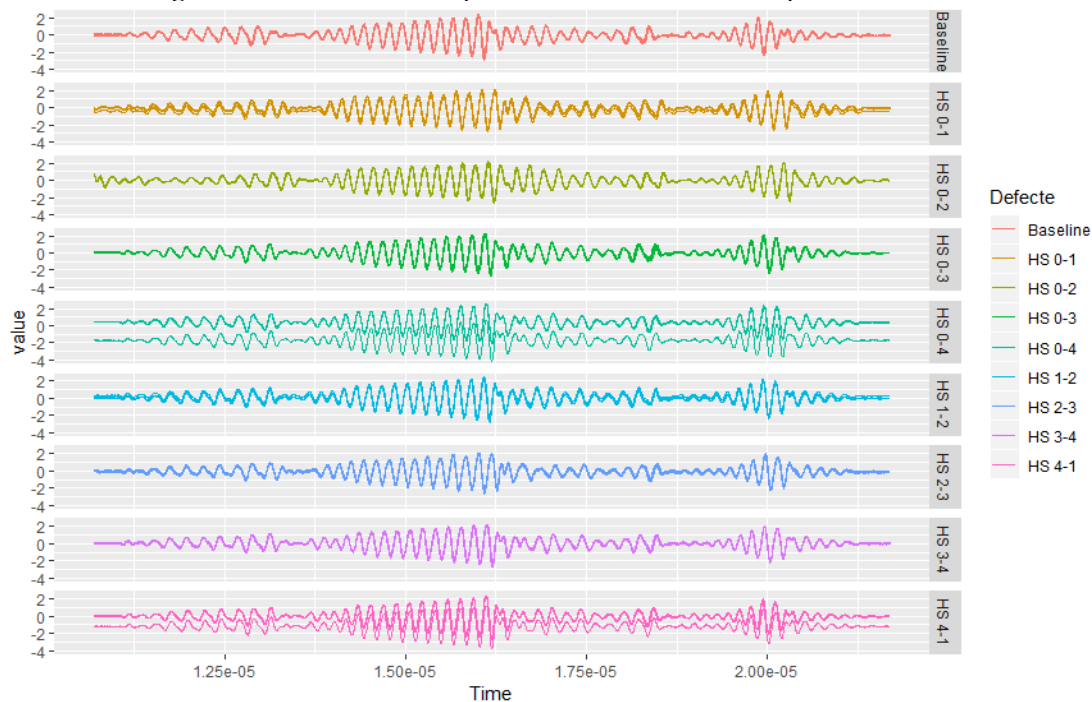


Figura 5.6: Señales de los experimentos de todos los casos para el sensor 4.

Detección de puntos calientes en estructuras utilizando ondas guiadas

Todas las señales tienen cierto grado de similitud, su amplitud siempre oscila en un rango de valores entre -4 y 2, excepto en algún caso que oscilan en un rango algo más amplio. Otra similitud que podemos observar es su forma, es semejante en todos los casos, aunque sí que se pueden apreciar ciertas diferencias.

No se puede concluir nada después de realizar este análisis preliminar. Observando las señales graficadas, se ve que son muy similares unas a otras, pero existen diferencias. No sabemos si dichas diferencias son relevantes o no, ya que no se dispone de más información. Para poder sacar conclusiones claras se debe realizar un estudio estadístico, como puede ser un análisis de componentes principales.

6.Detección y localización de puntos calientes

Una vez realizado el análisis preliminar, aún no se ha podido concluir nada relevante, ya que las diferencias observadas en las señales han sido mínimas. Por eso mismo se debe realizar un estudio más exhaustivo. Para ello, puesto que se dispone de muchos datos, se ha realizado un análisis de componentes principales con un número de componentes principales de 5. Para realizar el análisis se han utilizado dos métodos diferentes de cálculo (NIPALS y SVD) y dos tipos de escalado (AS y GS), obteniendo así 4 combinaciones diferentes: método NIPALS con AS, método SVD con AS, método NIPALS con GS y método SVD con GS. Además, antes de ejecutar el análisis se ha realizado una normalización previa de los datos para eliminar un offset que tenían algunas señales, por lo que también se compararán los resultados habiendo realizado este pre-escalado y sin haberlo realizado.

En resumen:

- Se han realizado los cálculos con dos tipos de escalado: Autoscaling y Groupscaling.
- Se han utilizado dos métodos de cálculo: SVD y NIPALS.
- Se han realizado los cálculos habiendo realizado una normalización preliminar y sin ella.

Como ya sabemos, para realizar PCA se debe crear un modelo de la estructura sana y proyectar en él los datos de los casos con defecto para poder compararlos entre ellos. Para crear dicho modelo se han utilizado 120 experimentos de los 150 que se realizaron en total para el caso baseline. Los demás experimentos, 30 para cada defecto y los 30 restantes del baseline, se han utilizado para la validación.

6.1. Resultados

En este apartado se mostrarán los resultados obtenidos a lo largo del análisis para cada método y tipo de escalado y se realizará una comparativa entre ellos. Se analizarán los resultados para los scores y los estadísticos T^2 y Q , ya que estos nos muestran si hay indicios de defectos o no en la estructura.

6.1.1.Scores

Habiendo realizado escalado previo al PCA

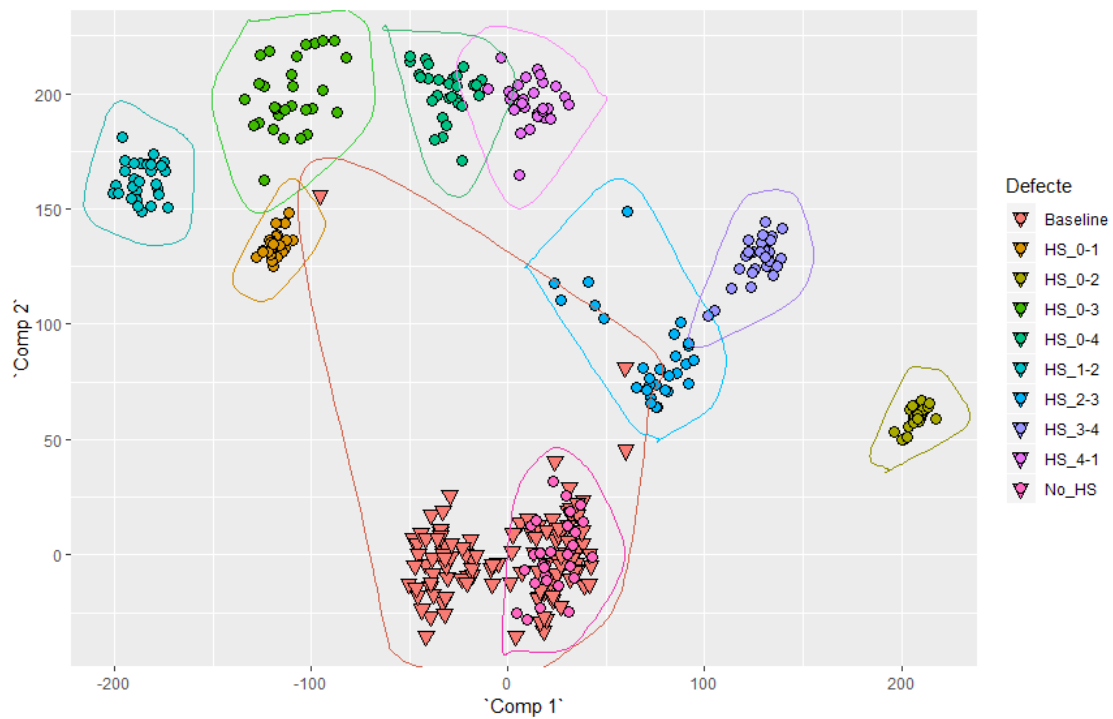


Figura 6.1: Proyección de todos los experimentos (con defecto y sin defecto) en el modelo PCA: score 1 vs score 2. Escalado previo. Método SVD. Escalado AS.

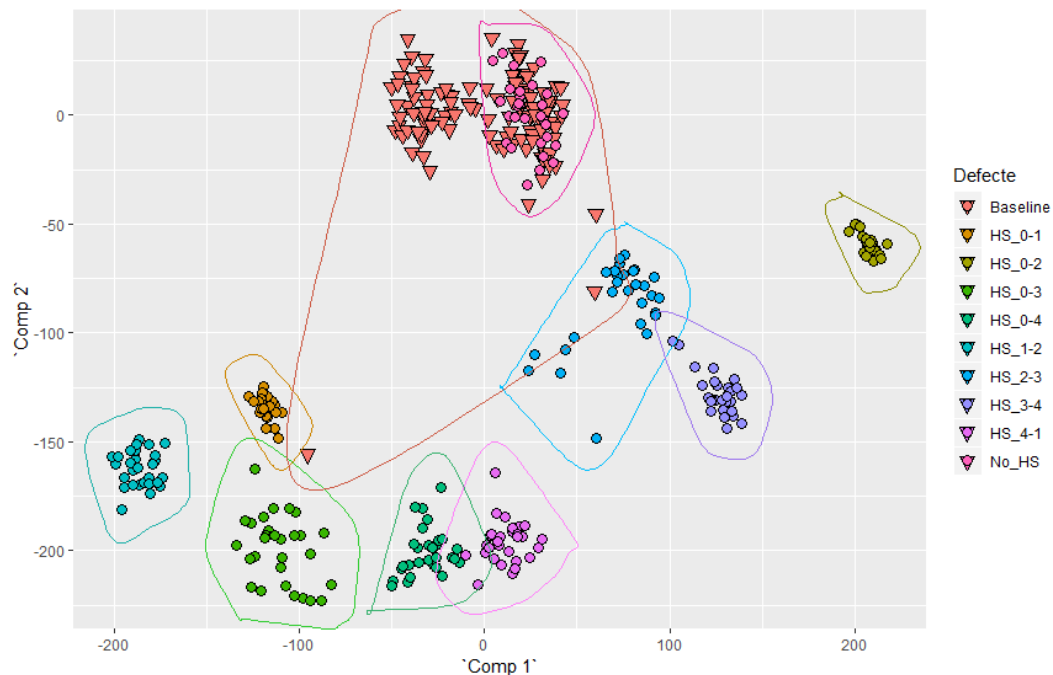


Figura 6.2: Proyección de todos los experimentos (con defecto y sin defecto) en el modelo PCA: score 1 vs score 2. Escalado previo. Método NIPALS. Escalado AS.

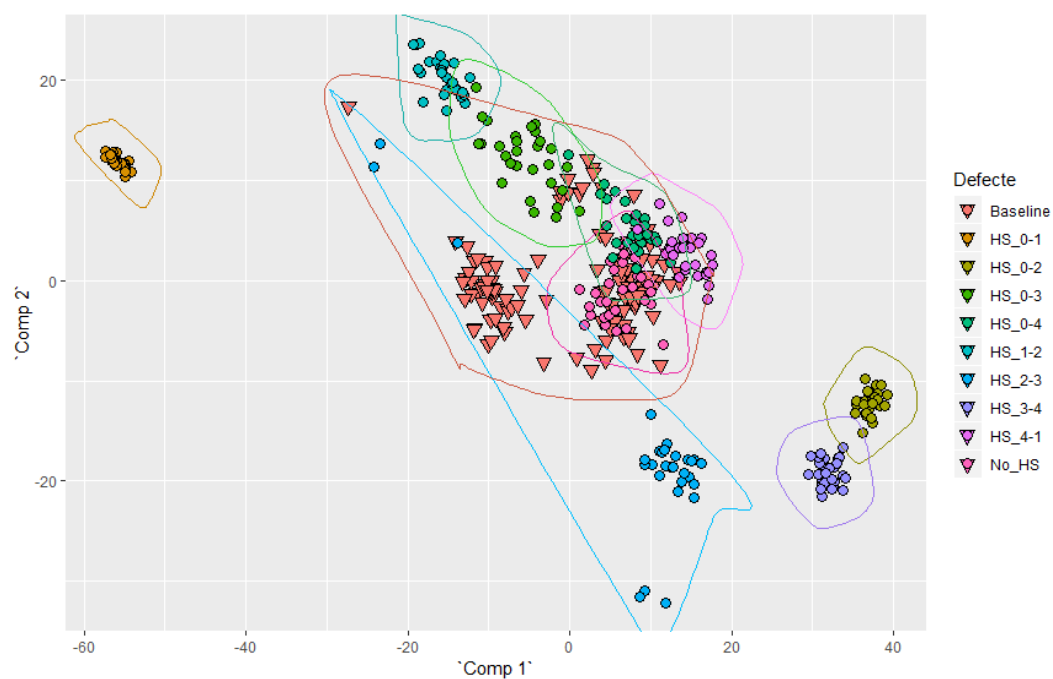


Figura 6.3: Proyección de todos los experimentos (con defecto y sin defecto) en el modelo PCA: score 1 vs score 2. Escalado previo. Método SVD. Escalado GS.

Detección de puntos calientes en estructuras utilizando ondas guiadas

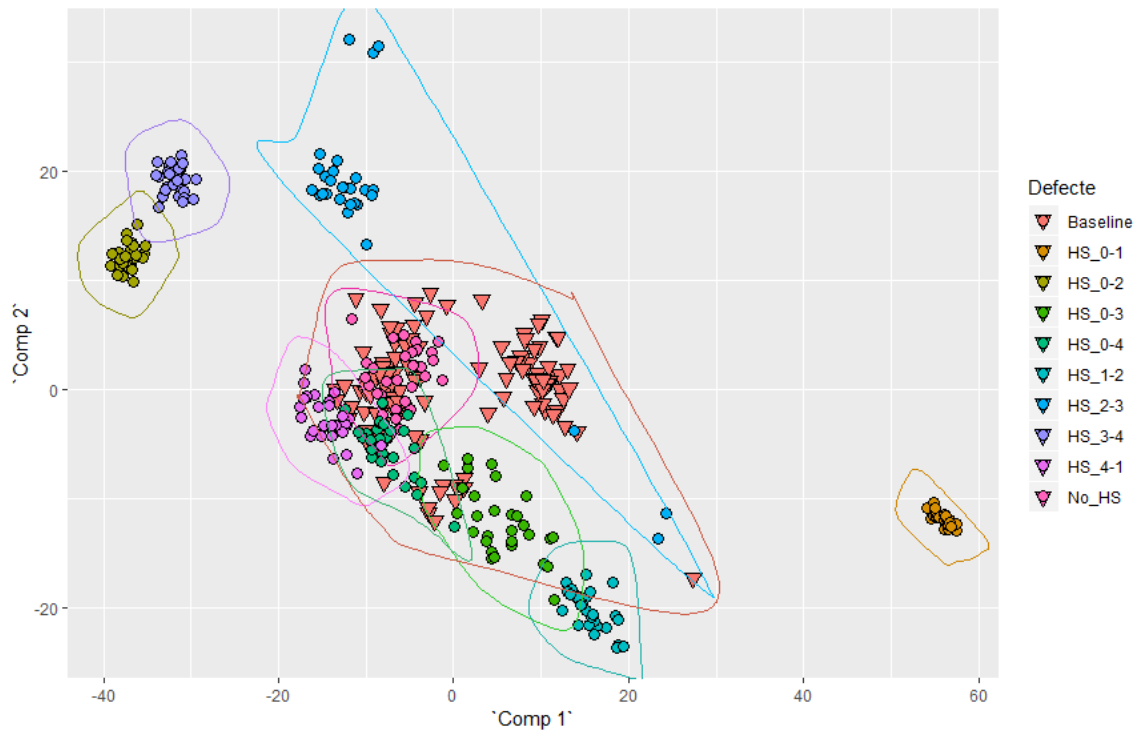


Figura 6.4: Proyección de todos los experimentos (con defecto y sin defecto) en el modelo PCA: score 1 vs score 2. Escalado previo. Método NIPALS. Escalado GS.

Si nos fijamos en la Figuras 6.1 y la Figura 6.2, en las que el tipo de escalado es el mismo y tan solo cambia el método de cálculo (SVD para la Figura 6.1 y NIPALS para la Figura 6.2), podemos observar que el único cambio entre ellas es el signo de los scores. Es decir, los scores que eran positivos para el cálculo con el método svd, pasan a ser negativos cuando el cálculo se realiza con el algoritmo NIPALS y viceversa. Lo mismo pasa para la Figura 6.3 y la Figura 6.4, el signo de los scores cambia y su valor absoluto se mantiene.

Si comparamos los gráficos teniendo en cuenta el tipo de escalado, se puede observar una gran diferencia entre AS y GS. Fijándonos en los gráficos donde el escalado utilizado ha sido AS (Figura 6.1 y Figura 6.2) podemos ver claramente que se distinguen 10 grupos de puntos, es decir, se distinguen los diferentes defectos en la estructura y también los experimentos del modelo baseline. Además, si nos

fijamos en los experimentos alrededor del origen (0,0), se puede observar como los datos sin defecto utilizados para la validación (No_HS), coinciden con los datos del modelo (baseline). En cambio, en la Figura 6.3 y la Figura 6.4, donde el escalado utilizado ha sido GS, cuesta más distinguir unos defectos de otros.

Sin haber realizado escalado previo al PCA

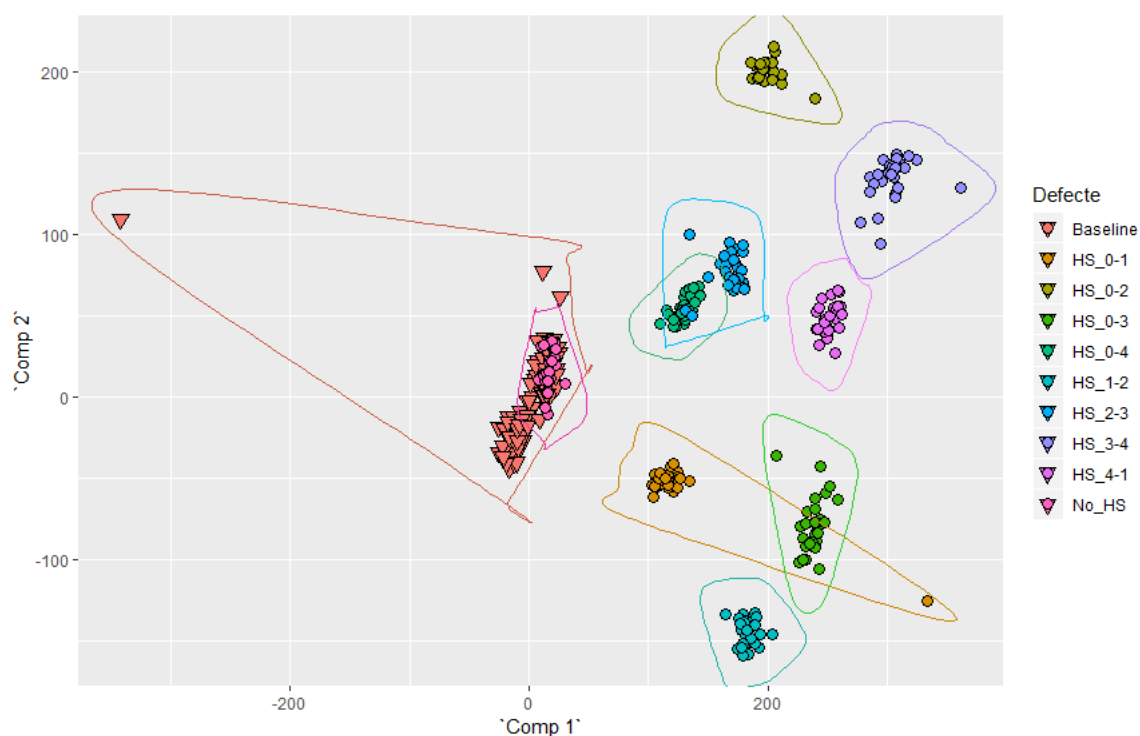


Figura 6.5: Proyección de todos los experimentos (con defecto y sin defecto) en el modelo PCA: score 1 vs score 2. Sin escalado previo. Método SVD. Escalado AS.

Detección de puntos calientes en estructuras utilizando ondas guiadas

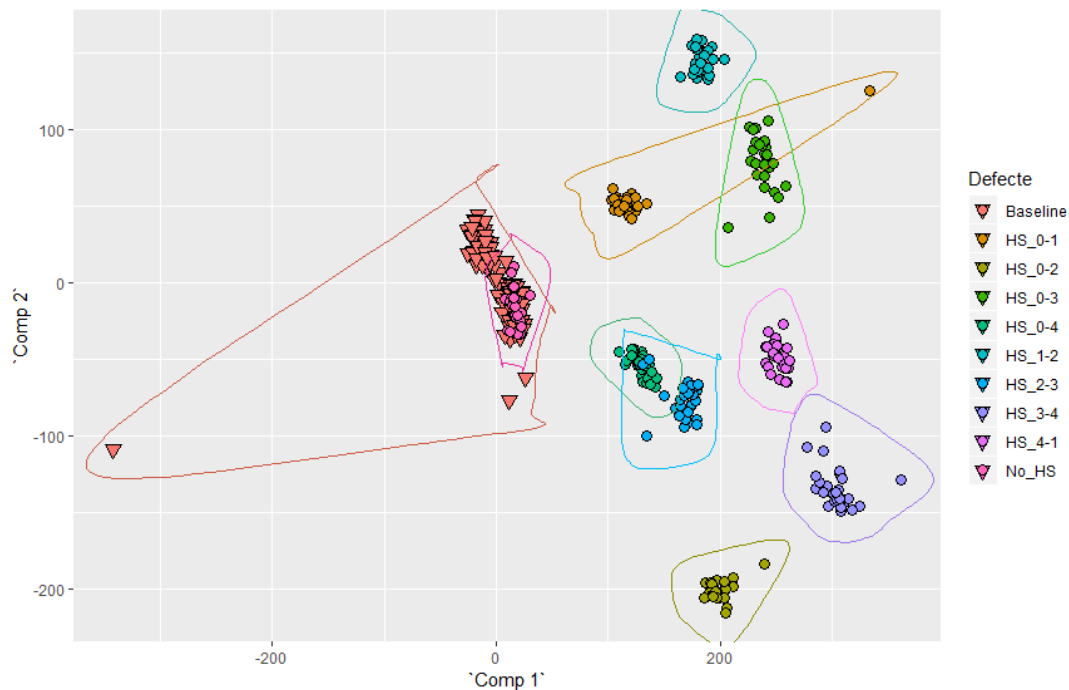


Figura 6.6: Proyección de todos los experimentos (con defecto y sin defecto) en el modelo PCA: score 1 vs score 2. Sin escalado previo. Método NIPALS. Escalado AS.

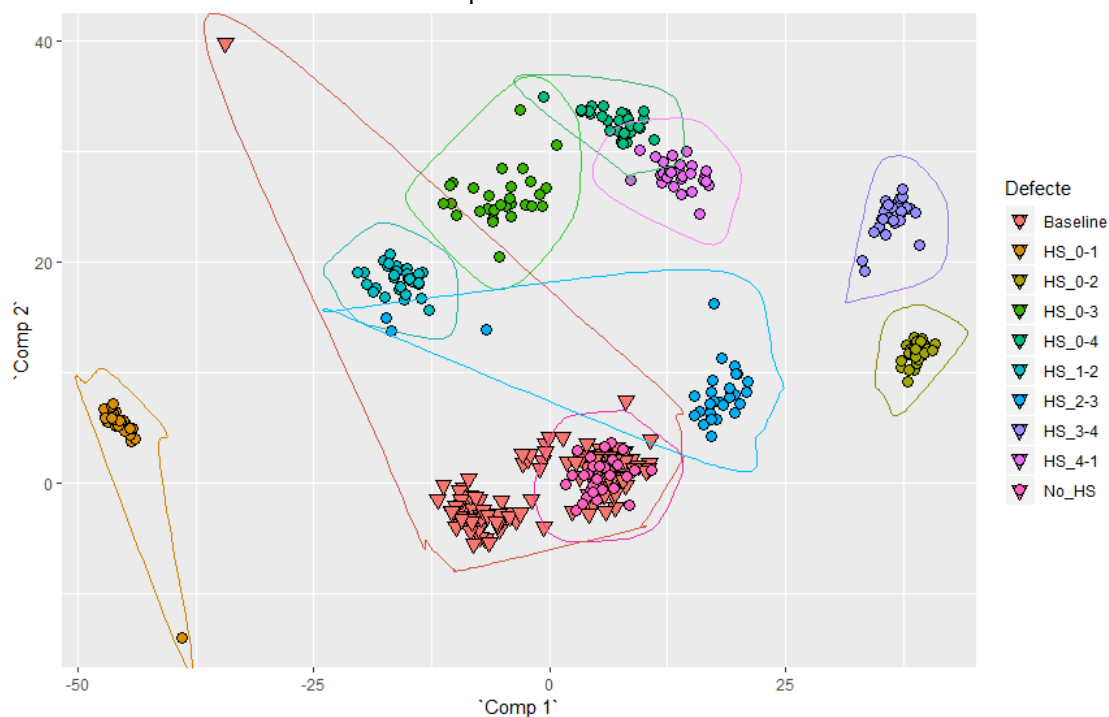


Figura 6.7: Proyección de todos los experimentos (con defecto y sin defecto) en el modelo PCA: score 1 vs score 2. Sin escalado previo. Método SVD. Escalado GS.

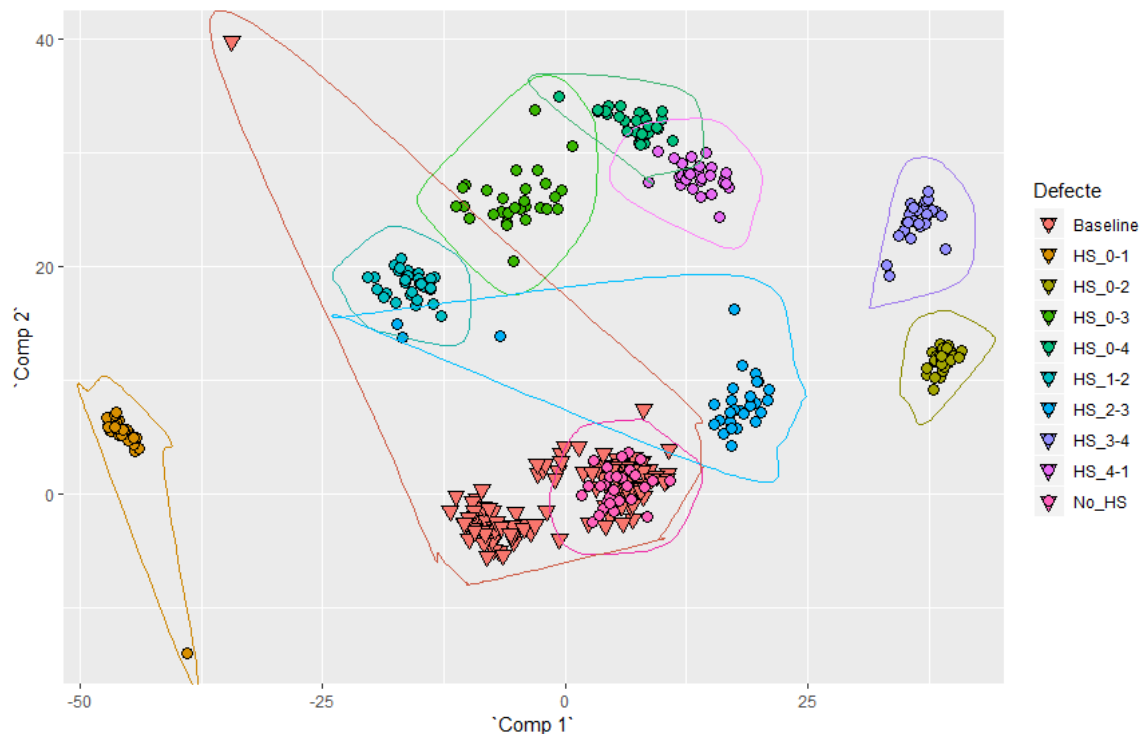


Figura 6.8: Proyección de todos los experimentos (con defecto y sin defecto) en el modelo PCA: score 1 vs score 2. Sin escalado previo. Método Nipals. Escalado GS.

Si nos fijamos en la Figura 6.5 y la Figura 6.6, donde se ha utilizado AS, ocurre lo mismo que se ha comentado anteriormente, los scores cambian de signo al aplicar SVD o NIPALS. También se puede apreciar cómo, exceptuando algún outlier, los grupos de puntos son bastante compactos y se pueden distinguir perfectamente todos los defectos.

En lo que se refiere a la Figura 6.7 y la Figura 6.8, donde se ha utilizado GS, ambos gráficos son idénticos, tanto para SVD como para NIPALS. Los defectos se pueden diferenciar unos de otros sin problema alguno aún con la existencia de outliers.

La existencia de dichos outliers que antes no aparecían, se debe al no haber realizado un escalado previo al cálculo de PCA para eliminarlos.

6.1.2. Estadístico T^2

Los estadísticos T^2 y Q solo varían según el tipo de escalado. Tanto para SVD como para NIPALS los gráficos son idénticos en caso de que el escalado aplicado sea el mismo, por lo que no se compararán los resultados según el método de cálculo utilizado.

El estadístico T^2 varía según el número de componentes principales con el que se haya realizado el análisis, contra mayor sea el número de componentes principales, mejores serán sus resultados.

Habiendo realizado escalado previo al PCA

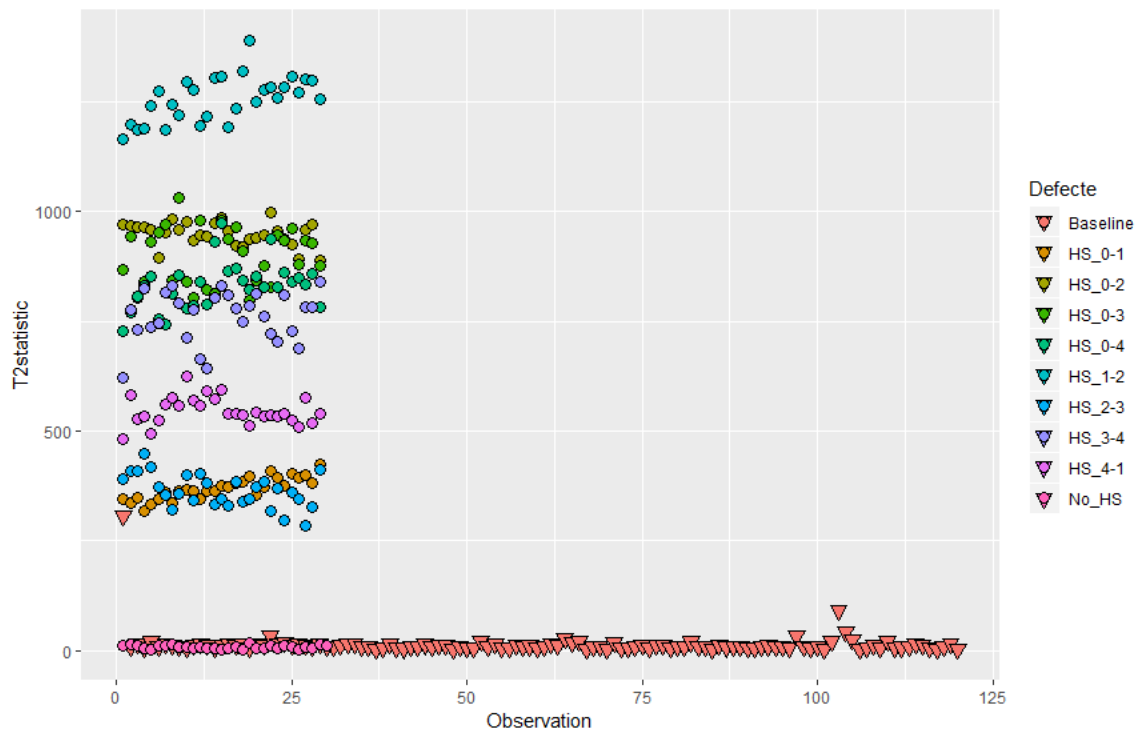


Figura 6.9: Proyección de todos los experimentos (con defecto y sin defecto) en el modelo PCA: estadístico T^2 . Escalado previo. Escalado AS.

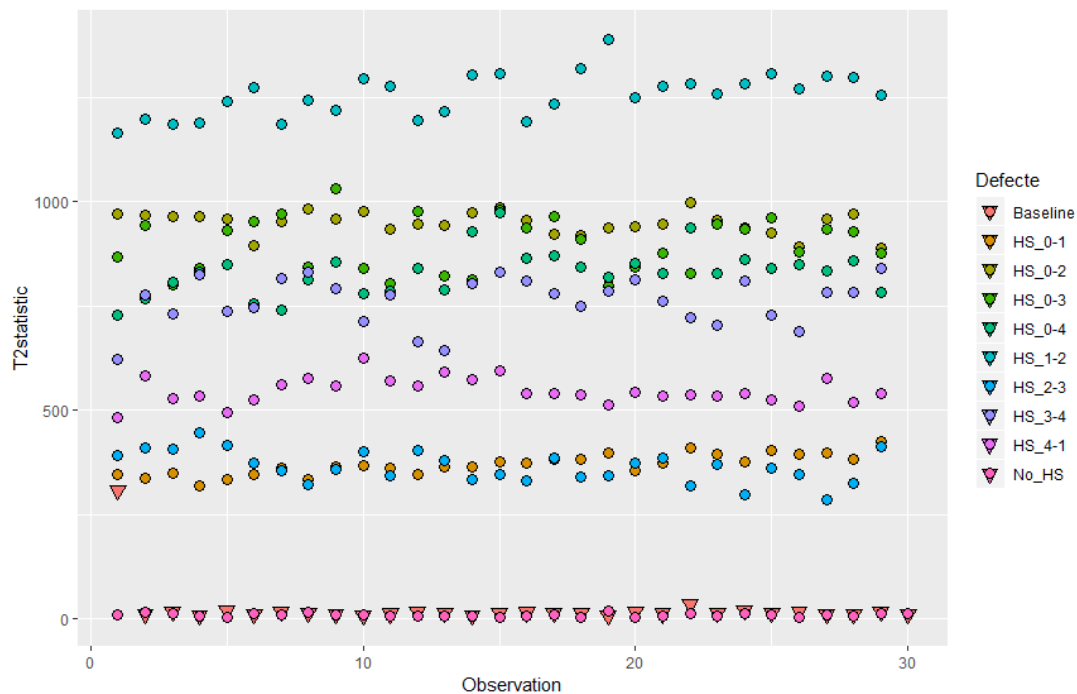


Figura 6.10: Proyección de los primeros 30 experimentos (con defecto y sin defecto) en el modelo PCA: estadístico T^2 . Escalado previo. Escalado AS.

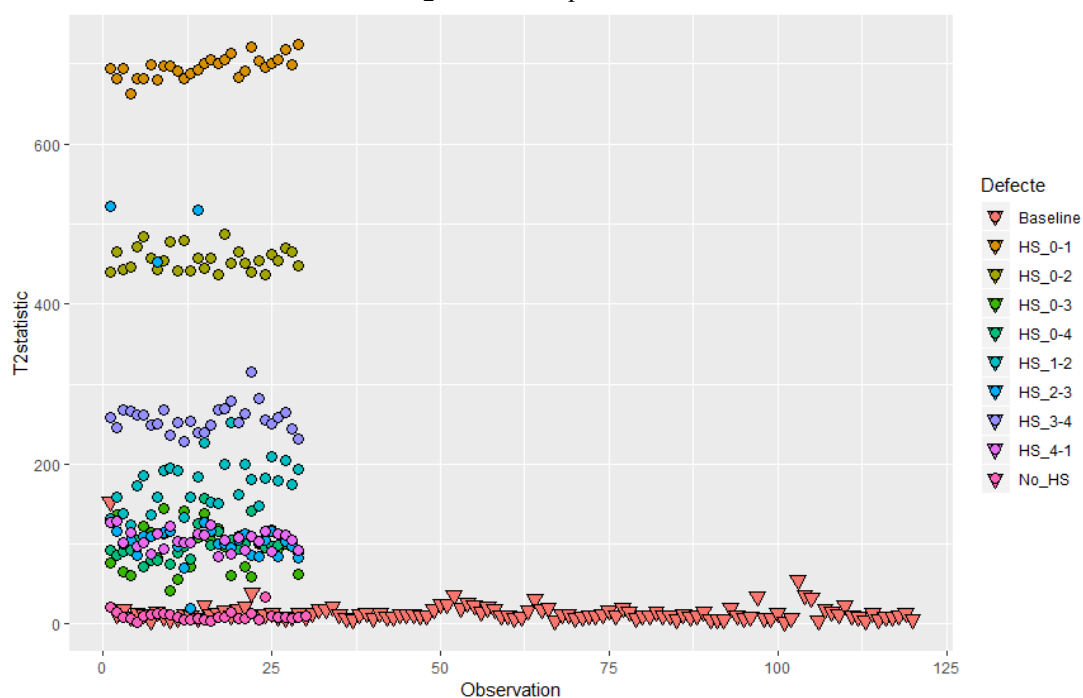


Figura 6.11: Proyección de todos los experimentos (con defecto y sin defecto) en el modelo PCA: estadístico T^2 . Escalado previo. Escalado GS.

Detección de puntos calientes en estructuras utilizando ondas guiadas

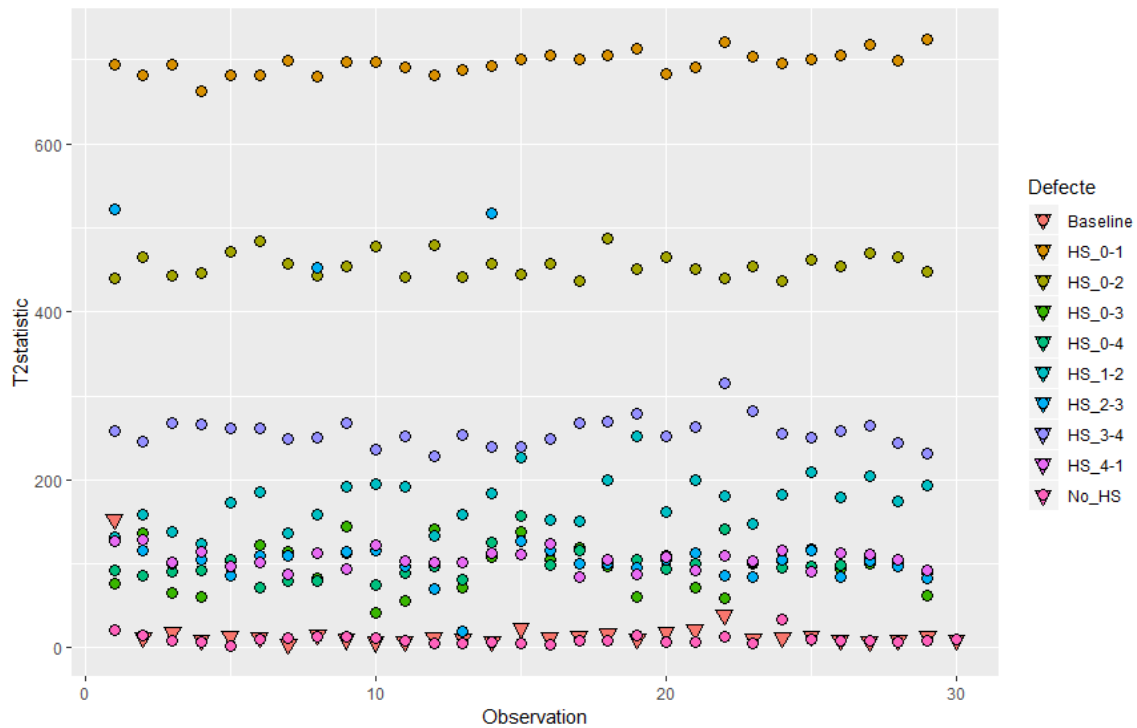


Figura 6.12: Proyección de los primeros 30 experimentos (con defecto y sin defecto) en el modelo PCA: estadístico T^2 . Escalado previo. Escalado GS.

Fijándonos en los gráficos se puede observar que los defectos no se pueden distinguir muy bien unos de otros tanto para AS (Figura 6.9 y Figura 6.10) como para GS (Figura 6.11 y Figura 6.12). Esto no importa demasiado, ya que puede ser que esa información nos la da el estadístico Q o los scores como ya hemos visto anteriormente.

Sin haber realizado escalado previo al PCA



Figura 6.13: Proyección de todos los experimentos (con defecto y sin defecto) en el modelo PCA: estadístico T^2 . Sin escalado previo. Escalado AS.

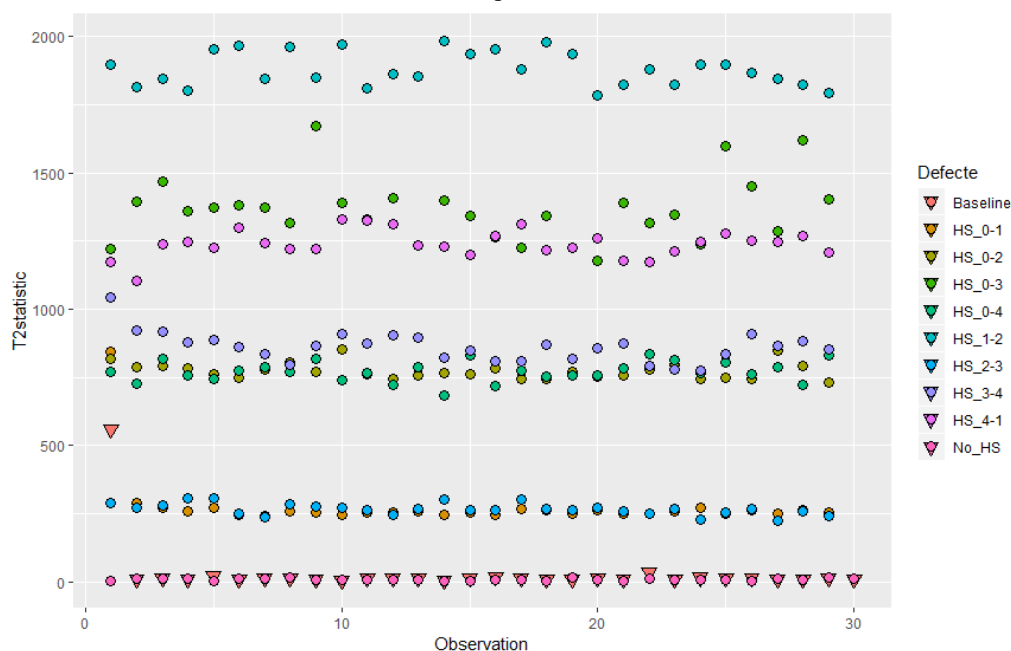


Figura 6.14: Proyección de los primeros 30 experimentos (con defecto y sin defecto) en el modelo PCA: estadístico T^2 . Sin escalado previo. Escalado AS.

Detección de puntos calientes en estructuras utilizando ondas guiadas

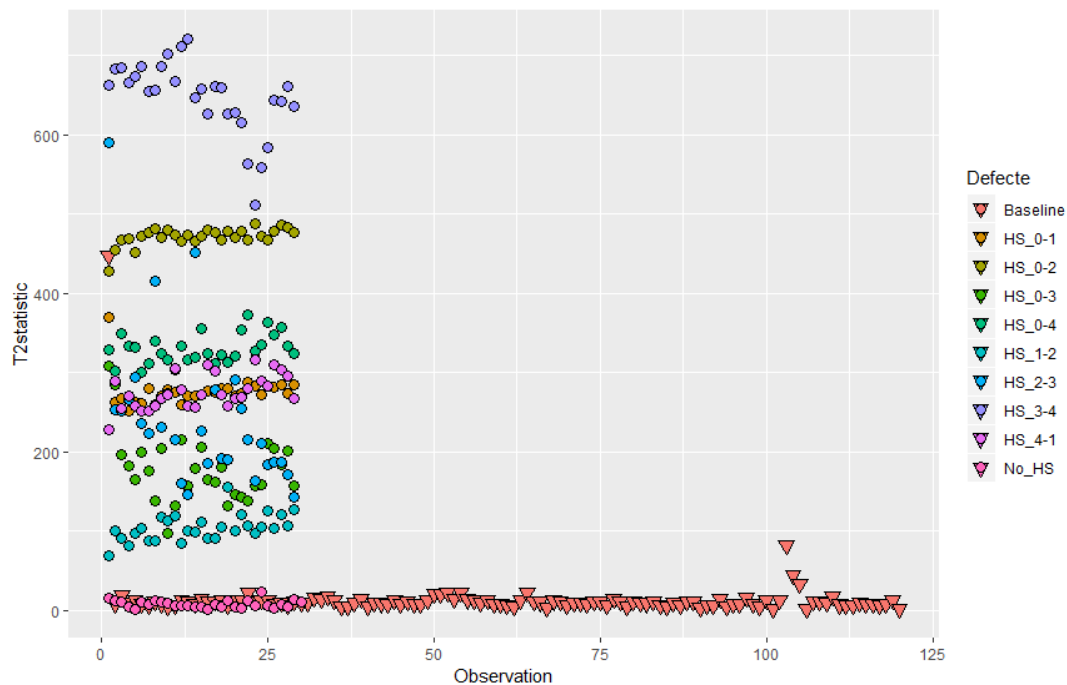


Figura 6.15: Proyección de todos los experimentos (con defecto y sin defecto) en el modelo PCA: estadístico T^2 . Sin escalado previo. Escalado GS.

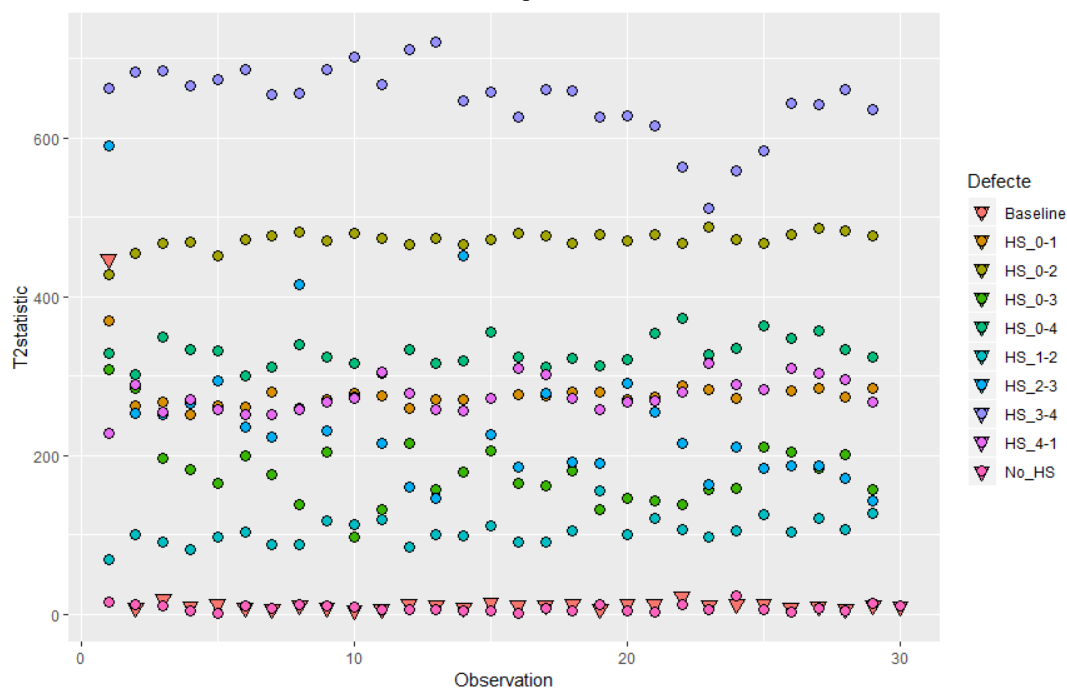


Figura 6.16: Proyección de los primeros 30 experimentos (con defecto y sin defecto) en el modelo PCA: estadístico T^2 . Sin escalado previo. Escalado GS.

Al igual que en los gráficos de T^2 donde se había realizado el escalado previo, podemos observar que si este no se realiza, los resultados no varían mucho. Los defectos no pueden distinguirse unos de otros, los puntos de cada caso están muy dispersos tanto para AS (Figura 6.13 y Figura 6.14) como para GS (Figura 6.15 y Figura 6.16).

6.1.3. Estadístico Q

Habiendo realizado escalado previo al PCA

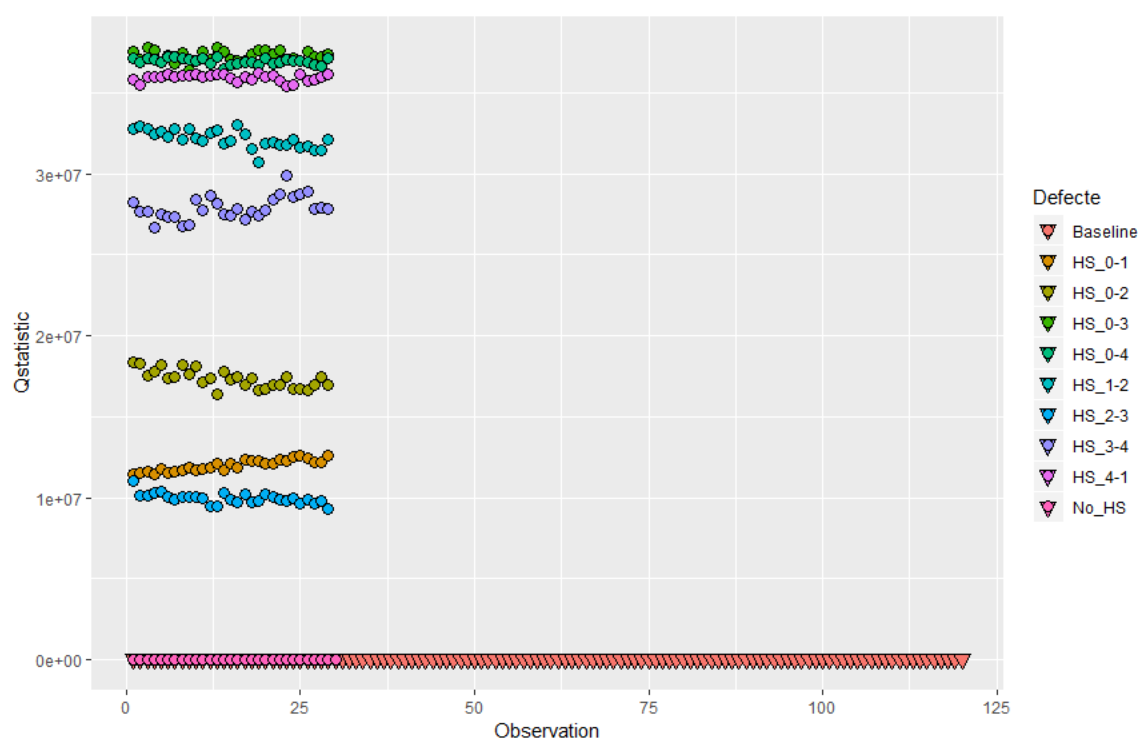


Figura 6.17: Proyección de todos los experimentos (con defecto y sin defecto) en el modelo PCA: estadístico Q. Escalado previo. Escalado AS.

Detección de puntos calientes en estructuras utilizando ondas guiadas

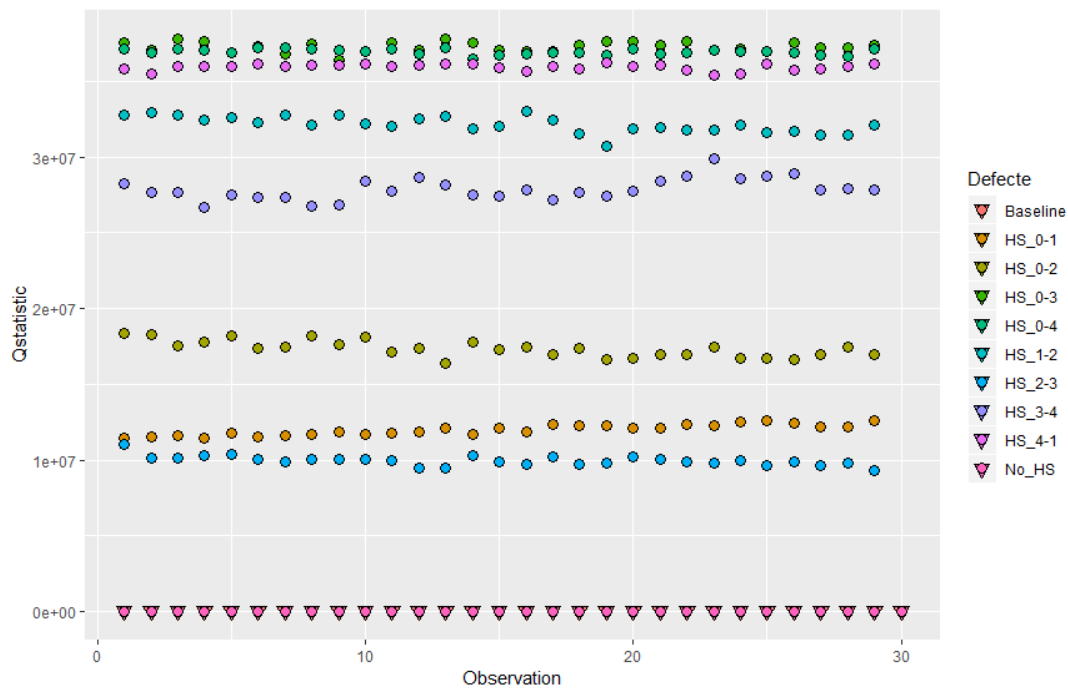


Figura 6.18: Proyección de los primeros 30 experimentos (con defecto y sin defecto) en el modelo PCA: estadístico Q. Escalado previo. Escalado AS.

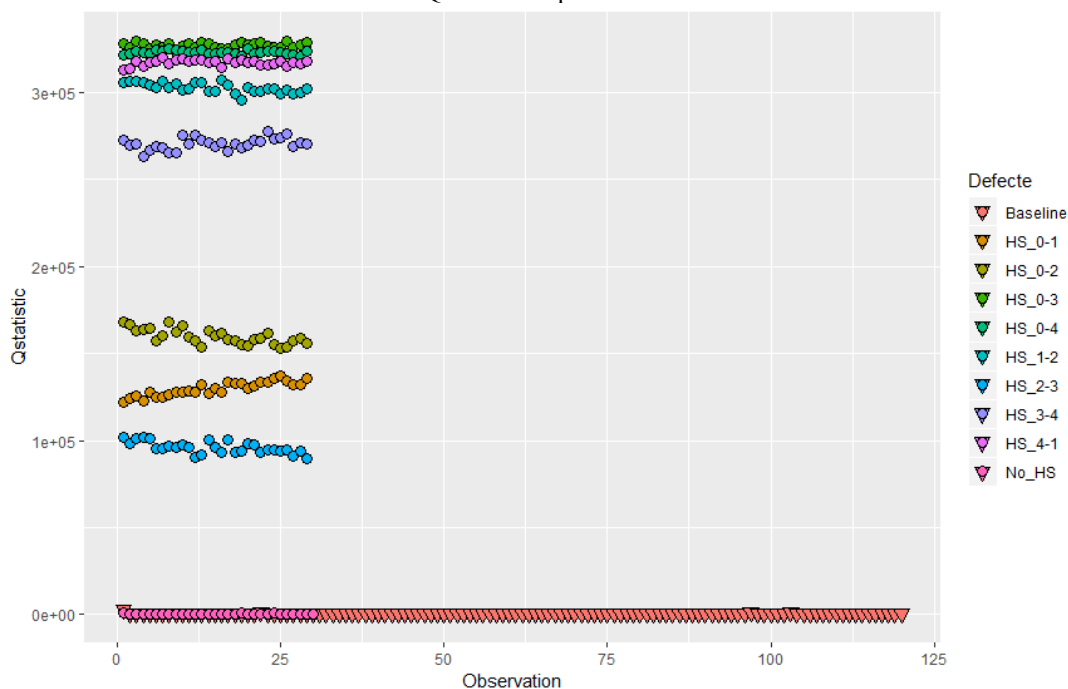


Figura 6.19: Proyección de todos los experimentos (con defecto y sin defecto) en el modelo PCA: estadístico Q. Escalado previo. Escalado GS.

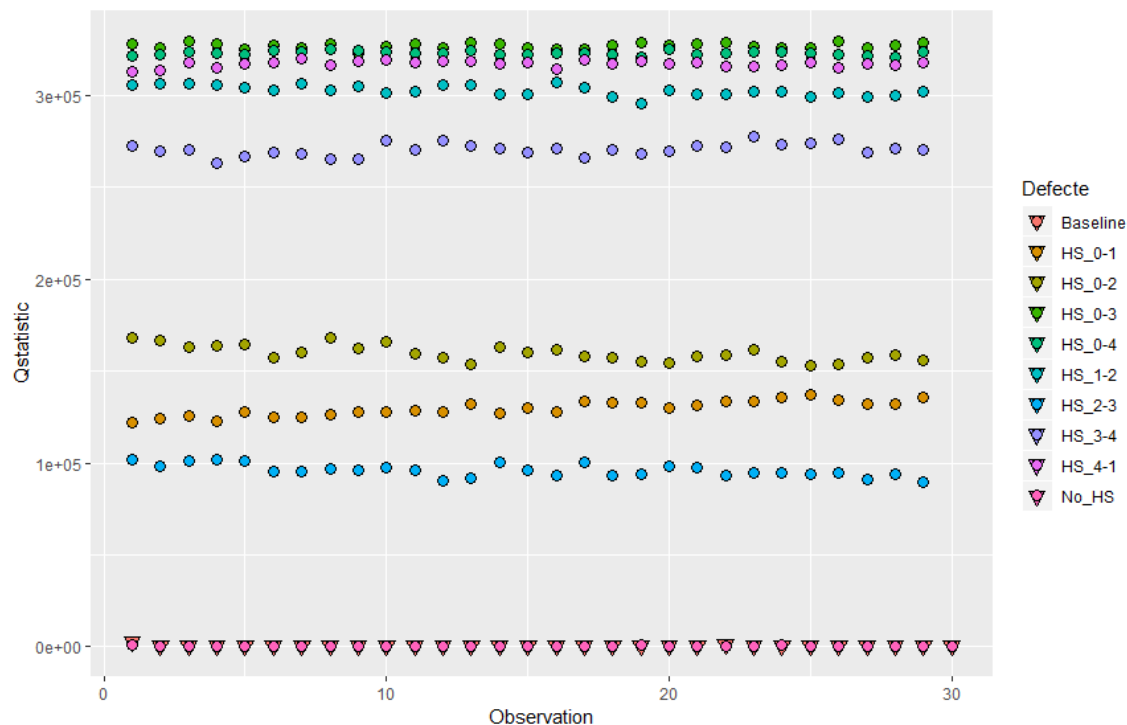


Figura 6.20: Proyección de los primeros 30 experimentos (con defecto y sin defecto) en el modelo PCA: estadístico Q. Escalado previo. Escalado GS.

Si comparamos los gráficos del estadístico Q, podemos observar que tanto para AS (Figura 6.17 y Figura 6.18) como para GS (Figura 6.19 y Figura 6.20) se pueden diferenciar la mayoría de los defectos sin problema alguno.

Sin haber realizado escalado previo al PCA

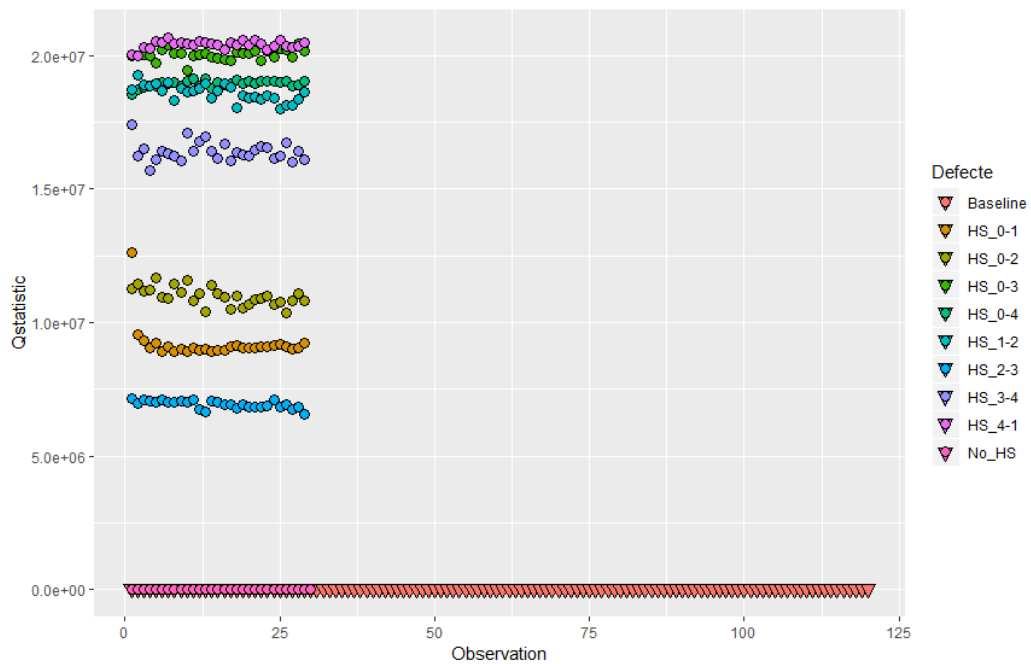


Figura 6.21: Proyección de todos los experimentos (con defecto y sin defecto) en el modelo PCA: estadístico Q. Sin escalado previo. Escalado AS.



Figura 6.22: Proyección de los primeros 30 experimentos (con defecto y sin defecto) en el modelo PCA: estadístico Q. Sin escalado previo. Escalado AS.

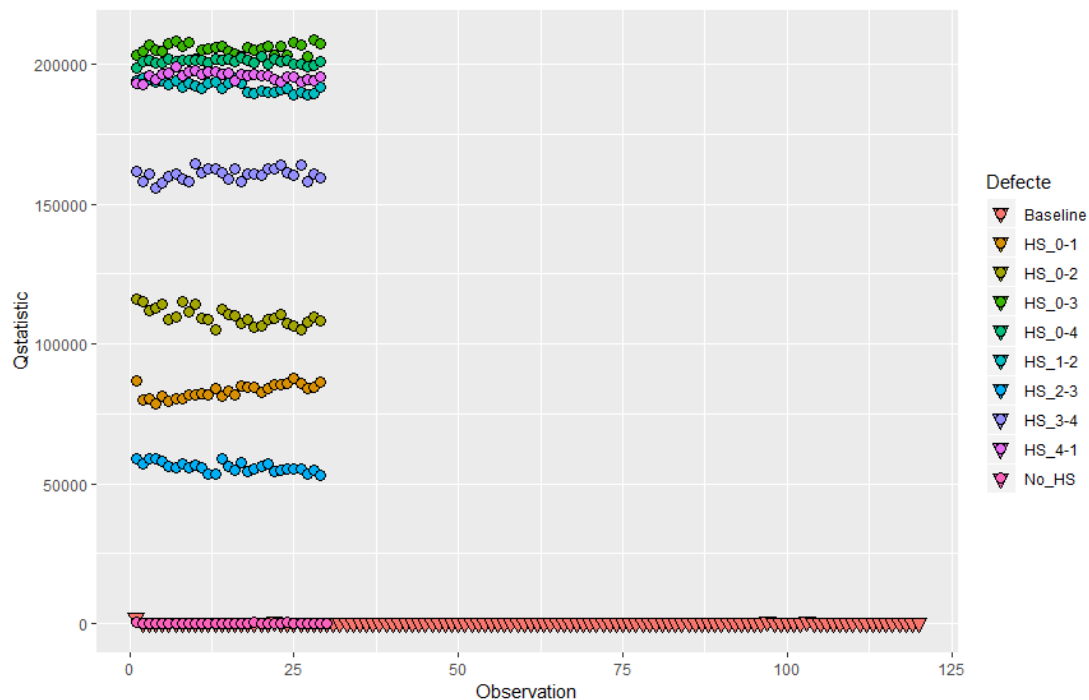


Figura 6.23: Proyección de todos los experimentos (con defecto y sin defecto) en el modelo PCA: estadístico Q. Sin escalado previo. Escalado GS.

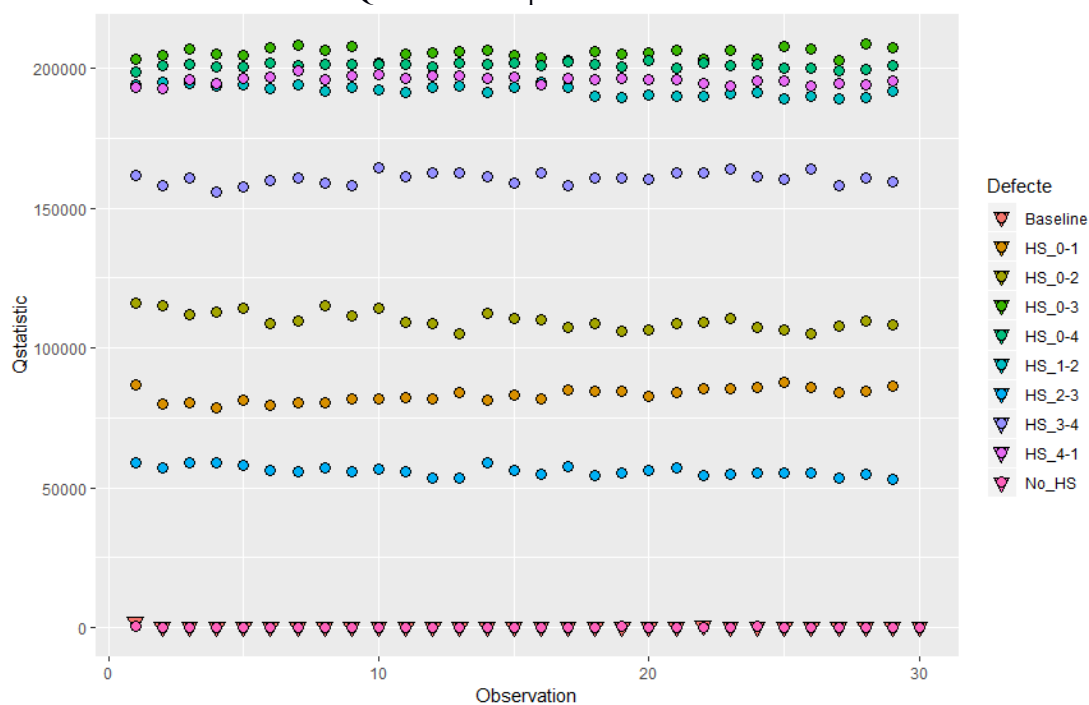


Figura 6.24: Proyección de los primeros 30 experimentos (con defecto y sin defecto) en el modelo PCA: estadístico Q. Sin escalado previo. Escalado GS.

Detección de puntos calientes en estructuras utilizando ondas guiadas

Al igual que pasaba con el estadístico Q cuando se realizaba el escalado previo, la mayoría de los defectos son distinguibles unos de otros. En este caso tampoco hay gran diferencia respecto a si se realiza el escalado previo o no ya los resultados son parecidos.

6.1.4. T^2 vs Q

Después de observar los gráficos para el estadístico T^2 y para el estadístico Q , se puede ver como Q discrimina mejor que T^2 . Como ya sabemos Q es más sensible que el estadístico T^2 debido a que este tiene una gran varianza y, por lo tanto, requiere de un gran cambio en las características de la estructura para que éste sea detectable.

En algunos casos graficar los scores es suficiente para distinguir defectos. En otros casos, el estadístico T^2 es más sensible a los cambios en la estructura, mientras que en otros el estadístico Q es mejor. Existen casos en que ni el estadístico T^2 ni el estadístico Q son lo suficientemente eficientes por separado, por eso es conveniente representar uno frente al otro en el mismo gráfico.

Habiendo realizado escalado previo al PCA

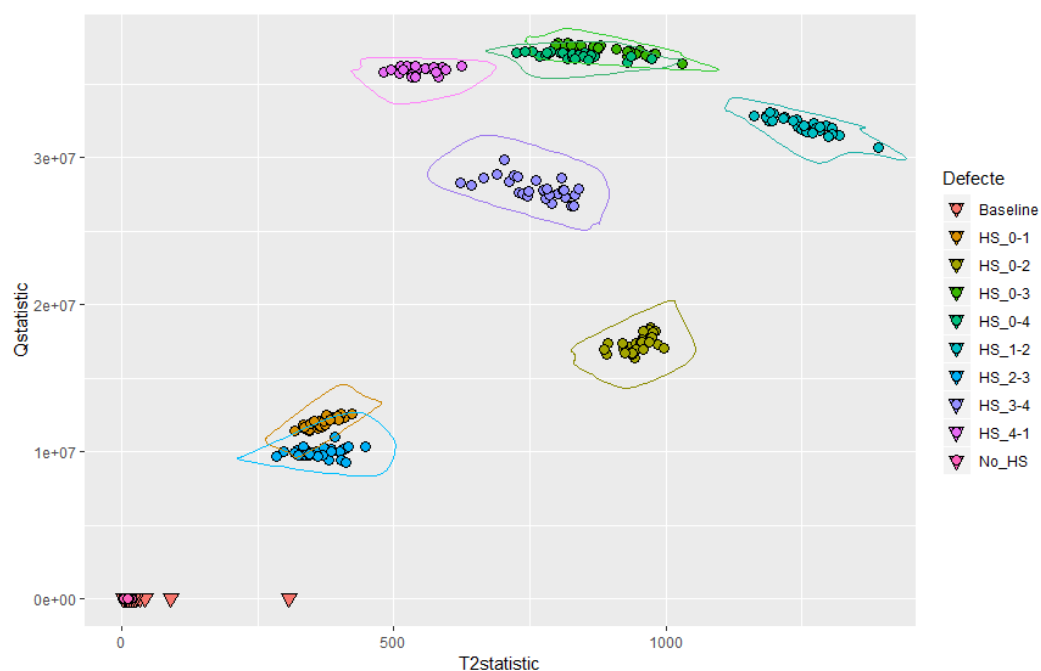


Figura 6.25: Proyección de todos los experimentos (con defecto y sin defecto) en el modelo PCA: T^2 vs Q. Escalado previo. Escalado AS.

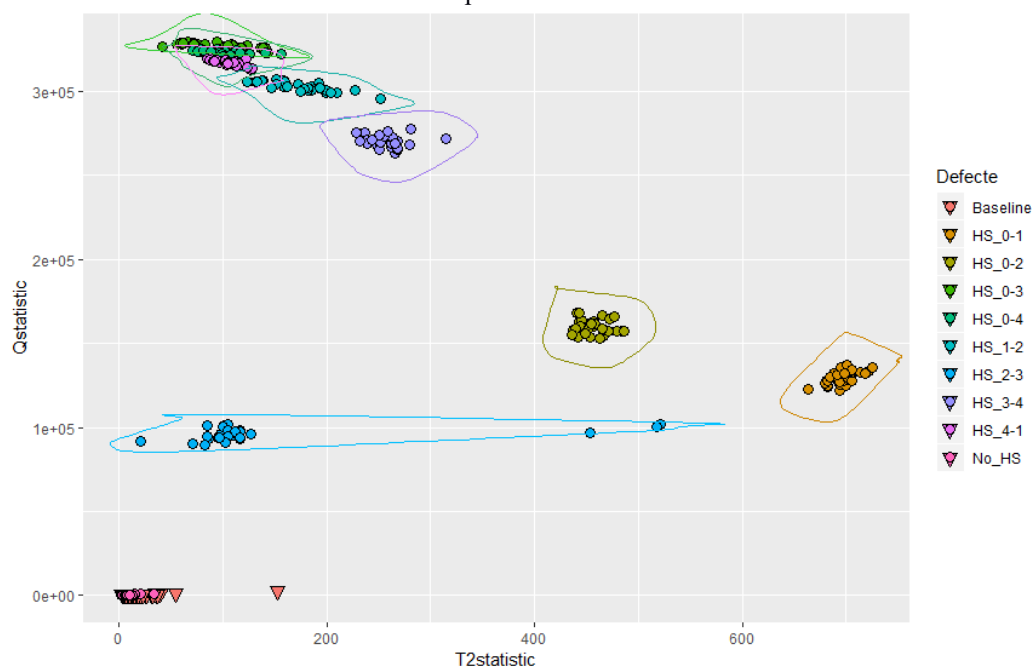


Figura 6.26: Proyección de todos los experimentos (con defecto y sin defecto) en el modelo PCA: T^2 vs Q. Escalado previo. Escalado GS.

Detección de puntos calientes en estructuras utilizando ondas guiadas

Al graficar los estadísticos T^2 y Q en un mismo gráfico se puede apreciar como la discriminación de los datos es mucho mejor que graficando cada estadístico por separado. En la figura 6.25, donde el método de escalado ha sido AS, se observa como todos los defectos son distinguibles entre ellos, exceptuando HS_0-3 y HS_0-4. En cambio en la Figura 6.26 hay varios grupos de puntos que están muy cercanos los unos de los otros y cuesta distinguirlos, aunque aún así se diferencian mejor que en el gráfico para el estadístico T^2 .

Sin haber realizado escalado previo al PCA

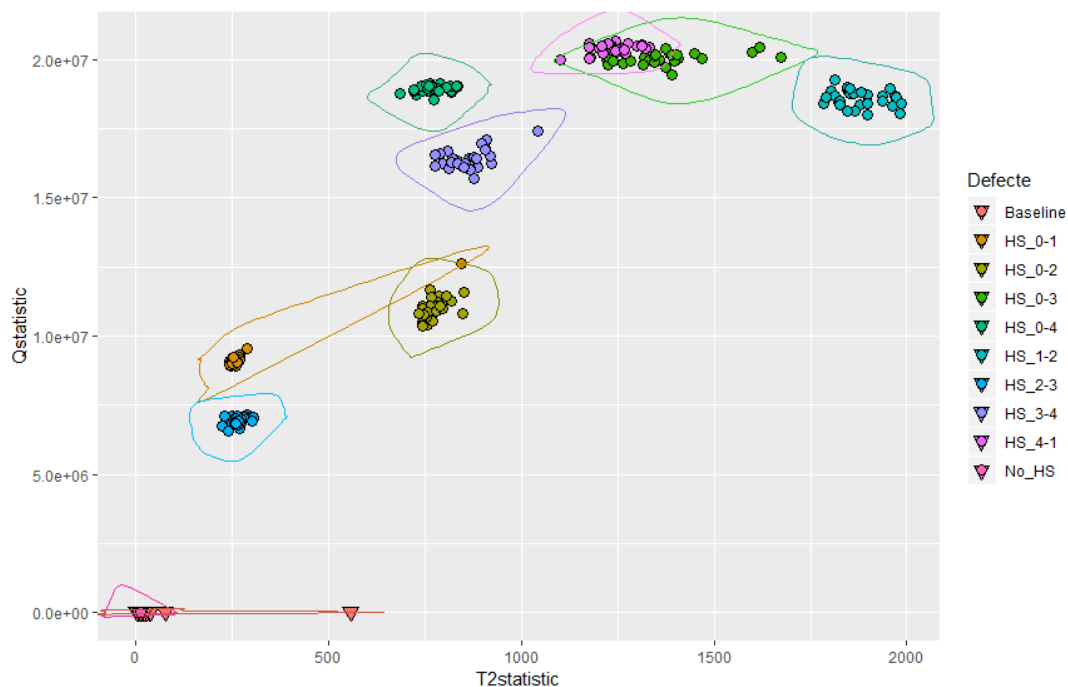


Figura 6.27: Proyección de todos los experimentos (con defecto y sin defecto) en el modelo PCA: T^2 vs Q. Sin escalado previo. Escalado AS.

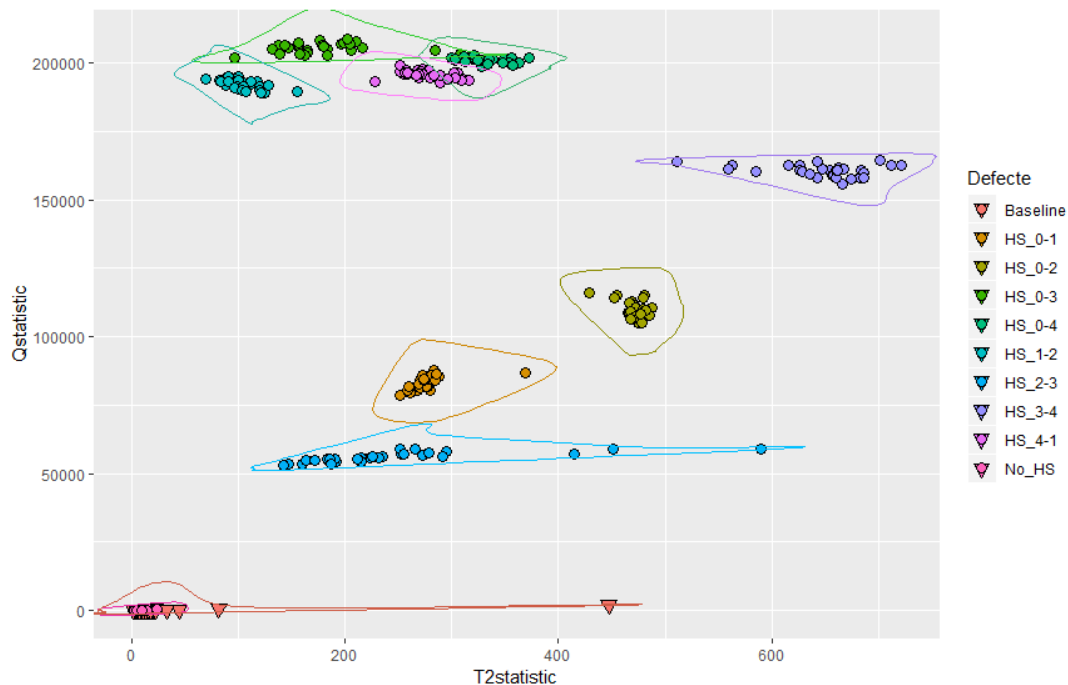


Figura 6.28: Proyección de todos los experimentos (con defecto y sin defecto) en el modelo PCA: T^2 vs Q . Sin escalado previo. Escalado GS.

Se puede observar como la discriminación de la figura 6.27 y la figura 6.28, donde no se ha realizado escalado previo, es mejor que en las figuras 6.25 y 6.26. En este caso, al contrario que en el anterior, en el gráfico que representa los datos en que el escalado utilizado ha sido GS (Figura 6.28), se pueden diferenciar algo mejor cada uno de los defectos.

7. Introducción a la localización de defectos

Una vez se ha determinado que la estructura tiene un defecto, se lleva a cabo la metodología para localizarlo. Para ello se analiza la contribución a cada sensor a los índices de detección de anomalías comentados en el apartado 4.2 del presente proyecto. Para calcular estas contribuciones se tiene en cuenta la energía y la desviación estándar de cada sensor y finalmente se grafican si es necesario.

Contribución del estadístico Q

El cálculo de esta contribución es bastante fácil, ya que es simplemente el error de una proyección elevado al cuadrado y sumado para todas las variables. La contribución de la variable jk para el estadístico Q del experimento i viene dada por [14]:

$$C(Q_i) = \tilde{x}_{i,jk}^2 = [\tilde{x}_i v_{jk}^T]^2 = [x_i (I - P P^T) v_{jk}^T]^2 \quad (35)$$

Donde v_{jk}^\square es un vector fila con JK elementos que nos indica la variable jk cuya contribución debe calcularse. Esto se hace estableciendo un valor de 1 a la posición del vector observación asociado con la variable de la cual se desea calcular su contribución, y el resto se establece en un valor de 0.

De esta forma, el vector de contribuciones para todas las variables es:

$$C(Q_i) = [\tilde{x}_{i,1}^2 \tilde{x}_{i,2}^2 \dots \tilde{x}_{i,JK}^2] = \left[(x_i (I - P P^T) v_1^T)^2 \dots (x_i (I - P P^T) v_{JK}^T)^2 \right] \quad (36)$$

Contribución del estadístico T^2

El estadístico T^2 depende de los scores y de las variables originales, en consecuencia, se puede calcular la influencia de cada variable original. Este estadístico se puede descomponer como el módulo al cuadrado de la contribución T^2 de la siguiente forma [14]:

$$T_i^2 = \|C(T_i^2)\|^2 = C(T_i^2)C(T_i^2)^T \quad (37)$$

Por otra parte, la ecuación (29) puede ser re-expresada como:

$$T_i^2 = x_i (P \wedge^{-1/2} \wedge^{-1/2} P^T) x_i^T \quad (38)$$

Relacionando las dos últimas ecuaciones, podemos determinar el vector de contribuciones de todas las variables:

$$C(T_i^2) = x_i P \wedge^{-1/2} = [C(T_i^2)_1 C(T_i^2)_2 \dots C(T_i^2)_{JK}] \quad (39)$$

Donde la contribución de las variables jk para el estadístico T^2 del experimento i viene dada por:

$$C(T_i^2)_{jk} = (x_i P \wedge^{-1/2} v_{jk}^T)^2 \quad (40)$$

7.1. Metodología para la localización de defectos

Para localizar defectos, se combinan los resultados del cálculo de las contribuciones para todas las fases de actuación. Basándonos en la hipótesis de que el defecto se localiza en la ruta entre el actuador y el sensor con mayor contribución, dicha ruta se marca como la región con mayor probabilidad de tener defecto. Sin embargo, los demás sensores también contribuyen, por lo tanto, existe una probabilidad de que el defecto esté localizado en alguna de estas rutas.

Detección de puntos calientes en estructuras utilizando ondas guiadas

De esta forma, se realiza un “mapa” de la superficie de la estructura indicando los actuadores, sensores y todas las rutas posibles. Las rutas son ponderadas según su contribución en cada fase. Finalmente, la superposición de estas rutas nos da la región con más probabilidad de la localización del defecto.

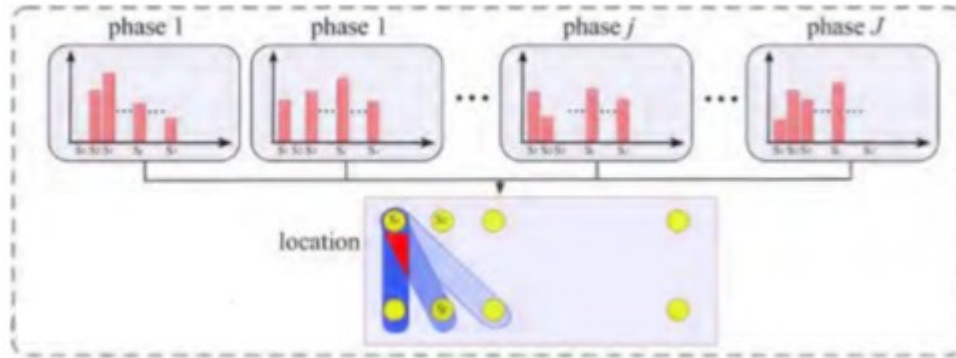


Figura 7.1: Localización de defectos [14].

7.2. Análisis preliminar de las contribuciones

Para calcular las contribuciones se ha utilizado el estadístico Q. El cálculo se ha realizado utilizando tanto los datos residuales como los loadings (matriz P) y el resultado ha sido idéntico, ya que estamos calculando la misma contribución pero de forma diferente.

A continuación encontramos las contribuciones calculadas con los datos a los que se les ha realizado un escalado previo al PCA, se ha utilizado el método svd y el tipo de escalado ha sido AS.

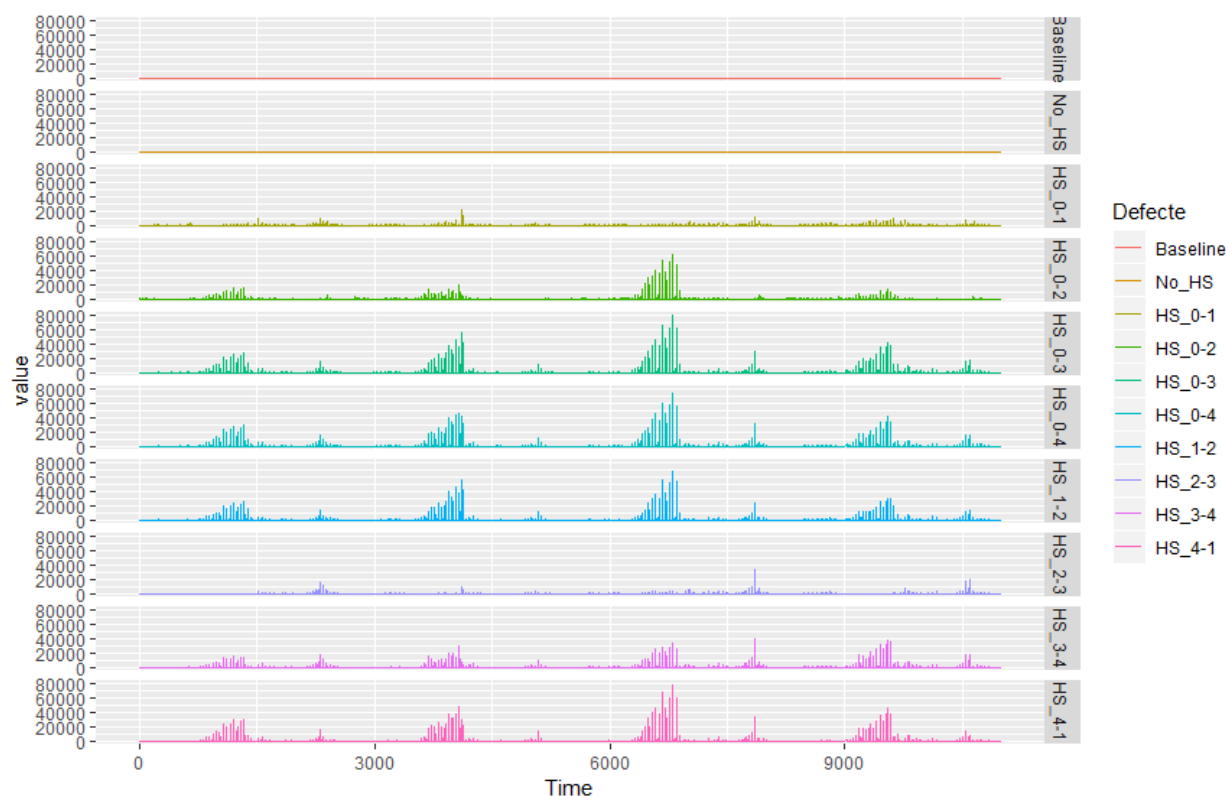


Figura 7.2: Ejemplo de contribuciones por sensores.

Como se puede observar, la figura 7.2 está organizada por los diferentes casos de estudio que se han realizado en los experimentos. Como ya sabemos, cada señal tiene 2750 muestras y tenemos 4 sensores, lo que hace un total de 11000 muestras si las graficamos para cada sensor una al lado de la otra. Las primeras 2750 muestras corresponden al sensor 1, las siguientes 2750 al sensor 2, etc.

Dicho esto, se puede apreciar como para el caso Baseline y el caso No_HS no hay contribución alguna por parte de ninguno de los sensores. Esto era de esperar, ya que no existe defecto en estos casos.

Para los demás casos, no se han obtenido los resultados que se esperaban. Para el caso HS_0-4 por ejemplo, se esperaba que el sensor con más contribución fuese el número 4, ya que el Hot Spot se encuentra entre él mismo y el actuador, pero no ha sido así. En ningún caso exceptuando baseline y

Detección de puntos calientes en estructuras utilizando ondas guiadas

No_HS se han obtenido los resultados esperados, de hecho, las contribuciones son muy parecidas en todos los casos y no debería ser así.



8. Análisis de impacto ambiental

Todos los equipos utilizados en los experimentos cumplen con la Directiva RoHS (Restriction of Hazardous Substances), abreviatura de la Directiva 2002/95/CE sobre la restricción del uso de ciertas sustancias peligrosas en equipos eléctricos y electrónicos.

También cumplen con la directiva WEEE (Waste Electrical and Electronic Equipment), abreviatura de la directiva 2002/96/CE que pretende promover el reciclaje, la reutilización y la recuperación de los residuos de los equipos eléctricos y electrónicos para reducir su contaminación.

El cableado utilizado está hecho con materiales permitidos por la Directiva RoHS y la soldadura se ha realizado con productos libres de Pb (plomo).

Además la Universitat Politècnica de Catalunya (UPC) sigue una política de eliminación de residuos con tal de minimizar el impacto ambiental.

Finalmente y no menos importante, la conservación y mantenimiento de todas las estructuras ayudan a prevenir cualquier tipo de desastre que pueda conllevar a la pérdida de vidas humanas y también a la contaminación del medio ambiente. Es verdad que el mismo establecimiento de una estructura en el planeta ya ha ocasionado una sacudida ambiental. Sin embargo, la mayoría de estas obras son levantadas para el mismo bienestar del hombre y de la vida en su entorno; por tanto, en caso que estas fallen, ocasionarán grandes colapsos que pueden llegar a destruir especies vegetales y animales (por ejemplo fallas en oleoductos, problemas con barcos cargueros, etc). Cada elemento o estructura que el hombre haya creado (incluidos todas aquellas estructuras con fines ecológicos, como por ejemplo los generadores eólicos), debe tener el correcto mantenimiento y vigilancia para preservar el planeta. Es así como este trabajo aporta un granito de arena más en la investigación para el mantenimiento de estructuras.

Conclusiones

El objetivo de este proyecto ha sido detectar puntos calientes en estructuras. Esto puede ser aplicable a muchos casos dentro de la ingeniería y resultar muy útil si se sigue desarrollando. Por esta razón, los resultados obtenidos en este proyecto serán útiles en un futuro dónde se quiera seguir avanzando con esta investigación. Ya que los puntos calientes se presentan en todas las estructuras que están sometidas a altos rendimientos (aviones), fricciones y vibraciones (obras civiles), altas temperaturas (hornos), etc., así como también los cambios bruscos de temperatura debidos a factores medio ambientales. Además, incluir este tipo de daño era de vital importancia en la investigación para el grupo CoDALab ya que hasta el momento no se había estudiado.

Si bien es cierto que la metodología implementada es completamente conocida en el área del SHM y por CoDALab, en este trabajo se parte de algo conocido e introducimos algo nuevo, en este caso un nuevo tipo de anomalía en la estructura. Es así como se pueden sacar varias conclusiones de los resultados obtenidos en este proyecto en lo que se refiere a las diferentes opciones que se han utilizado para realizar la detección de puntos calientes:

En relación al tipo de método de cálculo a la hora de realizar PCA, se ha observado que tanto SVD (Singular Value Decomposition) cómo el algoritmo NIPALS (Nonlinear Iterative Partial Least Square) han dado exactamente los mismos resultados. Es cierto que los scores cambian de signo al ejecutar un método u otro, pero eso no cambia la dificultad para distinguir unos defectos de otros en el gráfico.

Adicionalmente, se ha visto como el tipo de escalado influye bastante a la hora de obtener resultados. Normalmente groupscaling ofrece mejores resultados que autoscaling y así ha sido en algunos casos. Cuando se ha realizado un escalado previo al PCA, GS (groupscaling) ofrece peores resultados que AS (autoscaling). Se puede ver claramente en el gráfico de los scores que, para AS (Figura 6.1) los defectos son claramente distinguibles mientras que para GS (Figura 6.3) a penas se

pueden distinguir unos defectos de otros (exceptuando algún caso). En cambio, cuando no se ha realizado escalado previo al PCA, GS ofrece unos resultados donde, observando el gráfico podemos distinguir cada uno de los defectos. Una comparativa importante podría ser el gráfico de los scores para GS cuando se ha realizado escalado previo al PCA (Figura 6.3) y cuando dicho escalado no se ha realizado (Figura 6.7).

Pese a los diferentes caminos que se han seguido para realizar el estudio y obtener resultados, todos ofrecen algo en común, cumplen con el objetivo principal el proyecto, detectar puntos calientes. Ya sea con el gráfico de los scores, el del estadístico Q, el del estadístico T2o el gráfico de T2 contra Q, en alguno de estos casos, todos los caminos nos han mostrado una distinción clara de los defectos, llegando a la conclusión de que son detectables.

En lo que se refiere a la localización de puntos calientes, hemos visto que las contribuciones para el estadístico Q no mostraban lo que se esperaba. Por otra parte, no podemos sacar conclusiones claras, ya que simplemente se ha realizado un estudio preliminar.

Finalmente, todo el trabajo de montaje de los experimentos, el desarrollo de los algoritmos, el aprender sobre el tema de SHM, PCA y cómo trabajar con los datos almacenados han sido un gran reto pero a su vez me han dado una gran satisfacción personal que han valido la pena por el conocimiento porque me ha brindado un nuevo escenario científico que hasta entonces era desconocido para mí.

Presupuesto y/o análisis económica

MATERIAL			
CONCEPTO	COSTE POR UNIDAD	CANTIDAD	COSTE TOTAL
PZT	1,60 €/unidad	15 unidades	24,00 €
Adhesivo Araldite	8,95 €/unidad	1 unidad	8,95 €
Papel de lija	3,46 €/unidad	1 unidad	3,46 €
Cables	8 €/25metros	25 metros	8,00 €
Acetona	2,20 €/unidad	1 unidad	2,20 €
Pasta de soldadura	2,66 €/unidad	1 unidad	2,66 €
Estaño (unión para soldadura)	3,25 €/unidad	1 unidad	3,25 €
Guantes de latex	4,35 €/caja	1 caja	4,35 €

Tabla 1: Coste de material.

HARDWARE Y HERRAMIENTAS			
CONCEPTO	COSTE POR UNIDAD	CANTIDAD	COSTE TOTAL
Generador arbitrario de onda NI PXI-5412	158,60 €/unidad	1 unidad	158,00 €
Digitalizador de señales NI PXI-5114	89,00 €/unidad	1 unidad	89,00 €
Conmutador de matriz NI PXI-2529	86,20 €/unidad	1 unidad	86,20 €
Bloque terminal TB2636	11,20 €/unidad	1 unidad	11,20 €
Chasis con controlador integrado NI PXI-1033	42,30 €/unidad	1 unidad	42,30 €
Express card NI-8360	23,70 €/unidad	1 unidad	23,70 €
Ordenador Dell Latitude E-6530	41,70 €/unidad	1 unidad	41,70 €
Amplificador de señal Krohn-Hite 7602M	115,00 €/unidad	1 unidad	115,00 €
Soldador VTSS4N	1,00 €/unidad	1 unidad	1,00 €
Pistola de aire caliente Stanley FME670K	3,00 €/unidad	1 unidad	3,00 €
Cámara termográfica PCE-TC 3	330,00 €/unidad	1 unidad	330,00 €
Trípode Mnafrotto 055PROB	19,50 €/unidad	1 unidad	19,50 €
Termómetro con puntero láser AMPROBE IR608A	2,50 €/unidad	1 unidad	2,50 €
Osciloscopio PicoScope 5000	87,00 €/unidad	1 unidad	87,00 €

Tabla 2: Coste de hardware y herramientas.

COSTES DE INVESTIGACIÓN			
CONCEPTO	COSTE POR UNIDAD	CANTIDAD	COSTE TOTAL
Ingeniero: Jesús Romero	20,00 €/hora	620h	12.400,00 €
Director: Luis Eduardo Mujica	50,00 €/hora	120h	6.000,00 €
Codirectora: Magda Ruiz	50,00 €/hora	50h	2.500,00 €

Tabla 3: Costes de investigación.

COSTE TOTAL	
Material	56,87 €
Hardware y herramientas	1.010,10 €
Costes de investigación	20.900,00 €
TOTAL	21.966,97 €

Tabla 4: Coste total.

Los costes referentes a hardware y herramientas se han calculado suponiendo una vida útil de trabajo para cada dispositivo de cinco años y considerando que se han utilizado durante un tiempo de dos meses. Se ha calculado de esta forma puesto que estos dispositivos no son consumibles, es decir, no son de un solo uso.

Lista de tablas

Tabla 1: Coste de material.....	105
Tabla 2: Coste de hardware y herramientas.....	105
Tabla 3: Costes de investigación.....	106
Tabla 4: Coste total	106

Lista de figuras

Figura 1.1: Monitorización y detección de daños en estructuras - SHM [1].....	9
Figura 1.2: Metodología desarrollada en CoDALab para SHM [Propiedad del grupo CoDALab].....	10
Figura 1.3: Principio de operación de los piezoeléctricos (PZT's) [2].....	11
Figura 1.4: Detección de daños o defectos en estructuras [Propiedad del grupo CoDALab].....	11
Figura 2.1: Representación esquemática de las ondas longitudinales [5].....	16
Figura 2.2: Representación esquemática de las ondas transversales [5].....	16
Figura 2.3: Representación esquemática de las ondas de Rayleigh [5].....	16
Figura 2.4: Representación esquemática de las ondas de Rayleigh en sólidos isotrópicos [6].....	17
Figura 2.5: Representación esquemática de las ondas de Lamb [5].....	18
Figura 2.6: Comportamiento modo de vibración simétrico [5].....	19
Figura 2.7: Comportamiento modo de vibración anti-simétrico [5].....	19
Figura 2.8: Representación esquemática de las ondas de Stonely [5].....	20
Figura 2.9: Representación esquemática de las ondas de Love [7].....	20
Figura 2.10: Efecto de la temperatura en la amplitud de la señal [5].....	21
Figura 2.11: Efecto de la temperatura en el tiempo de propagación de la onda [5].....	22
Figura 3.1: Generador arbitrario de onda PXI-5412.....	24
Figura 3.2: Digitalizador de señales PXI-5114.....	25
Figura 3.3: Conmutador de matriz PXI-2529 y bloque terminal TB2636.....	26
Figura 3.4: Chasis PXI-1033.....	27
Figura 3.5: Express card NI-8360.....	27
Figura 3.6: Ordenador Dell Latitude E-6530 [8].....	28
Figura 3.7: Amplificador 7602M Wideband Amplifier [9].....	29
Figura 3.8: Soldador VTSS4N [10].....	30
Figura 3.9: Funcionamiento de los piezoeléctricos [Propiedad del grupo CoDALab].	31
Figura 3.10: Pistola de aire caliente FME670K.....	32
Figura 3.11: Cámara termográfica PCE-TC 3 [11].....	32

Figura 3.12: Termómetro con puntero laser IR608A.....	33
Figura 3.13: Estructura de estudio.....	34
Figura 3.14: PicoScope 5000 Series [12].....	35
Figura 3.15: Diagrama de conexión máximo dos salidas [12].....	35
Figura 3.16: Diagrama de conexión mínimo tres salidas [12].....	35
Figura 3.17: Diagrama de conexión PicoScope 5000 Series [12].....	36
Figura 3.18: Configuración inicial: posicionamiento de los 15 PZT's.....	38
Figura 3.19: Zona de estudio y posicionamiento de los PZT.....	40
Figura 3.20: Papel de lija.....	41
Figura 3.21: Pegamento Araldite Crystal.....	41
Figura 3.22: Conexionado PZT.....	42
Figura 3.23: Posicionamiento de la cámara termográfica respecto a la placa.....	43
Figura 3.24: Generación de puntos calientes y control de su temperatura.....	44
Figura 3.25: Señal utilizada en los experimentos.....	45
Figura 3.26: Diagrama de conexión entre componentes.....	46
Figura 3.27: Hot Spot 0-1.....	47
Figura 3.28: Hot Spot 0-2.....	47
Figura 3.29: Hot Spot 0-3.....	48
Figura 3.30: Hot Spot 0-4.....	48
Figura 3.31: Hot Spot 1-2.....	48
Figura 3.32: Hot Spot 2-3.....	48
Figura 3.33: Hot Spot 3-4.....	48
Figura 3.34: Hot Spot 4-1.....	48
Figura 4.1: Representación de los datos originales para el ejemplo de PCA.....	58
Figura 4.2: Datos del ejemplo PCA normalizados.....	59
Figura 4.3: Primer y segundo componente principal para el ejemplo de PCA.....	62
Figura 4.4: Despliegue de la matriz de datos inicial [15].....	67
Figura 4.5: Señales de los datos sin escalar de 4 sensores para el mismo experimento.....	68

Figura 4.6: Señales de los datos escalados de 4 sensores para el mismo experimento.....	69
Figura 5.1: Señales de los experimentos del caso baseline sobre el mismo eje de tiempo.....	71
Figura 5.2: Comparación señales caso baseline y caso HS 0-2 sobre el mismo eje de tiempo.....	72
Figura 5.3: Señales de los experimentos de todos los casos para el sensor 1.....	73
Figura 5.4: Señales de los experimentos de todos los casos para el sensor 2.....	73
Figura 5.5: Señales de los experimentos de todos los casos para el sensor 3.....	74
Figura 5.6: Señales de los experimentos de todos los casos para el sensor 4.....	74
Figura 6.1: Proyección de todos los experimentos (con defecto y sin defecto) en el modelo PCA: score 1 vs score 2. Escalado previo. Método SVD. Escalado AS.....	77
Figura 6.2: Proyección de todos los experimentos (con defecto y sin defecto) en el modelo PCA: score 1 vs score 2. Escalado previo. Método NIPALS. Escalado AS.	78
Figura 6.3: Proyección de todos los experimentos (con defecto y sin defecto) en el modelo PCA: score 1 vs score 2. Escalado previo. Método SVD. Escalado GS.....	78
Figura 6.4: Proyección de todos los experimentos (con defecto y sin defecto) en el modelo PCA: score 1 vs score 2. Escalado previo. Método NIPALS. Escalado GS.....	79
Figura 6.5: Proyección de todos los experimentos (con defecto y sin defecto) en el modelo PCA: score 1 vs score 2. Sin escalado previo. Método SVD. Escalado AS.....	80
Figura 6.6: Proyección de todos los experimentos (con defecto y sin defecto) en el modelo PCA: score 1 vs score 2. Sin escalado previo. Método NIPALS. Escalado AS.....	81
Figura 6.7: Proyección de todos los experimentos (con defecto y sin defecto) en el modelo PCA: score 1 vs score 2. Sin escalado previo. Método SVD. Escalado GS.....	81
Figura 6.8: Proyección de todos los experimentos (con defecto y sin defecto) en el modelo PCA: score 1 vs score 2. Sin escalado previo. Método Nipals. Escalado GS.....	82
Figura 6.9: Proyección de todos los experimentos (con defecto y sin defecto) en el modelo PCA: estadístico T^2 . Escalado previo. Escalado AS.....	83
Figura 6.10: Proyección de los primeros 30 experimentos (con defecto y sin defecto) en el modelo PCA: estadístico T^2 . Escalado previo. Escalado AS.....	84

Figura 6.11: Proyección de todos los experimentos (con defecto y sin defecto) en el modelo PCA: estadístico T^2 . Escalado previo. Escalado GS.....	84
Figura 6.12: Proyección de los primeros 30 experimentos (con defecto y sin defecto) en el modelo PCA: estadístico T^2 . Escalado previo. Escalado GS.....	85
Figura 6.13: Proyección de todos los experimentos (con defecto y sin defecto) en el modelo PCA: estadístico T^2 . Sin escalado previo. Escalado AS.....	86
Figura 6.14: Proyección de los primeros 30 experimentos (con defecto y sin defecto) en el modelo PCA: estadístico T^2 . Sin escalado previo. Escalado AS.....	86
Figura 6.15: Proyección de todos los experimentos (con defecto y sin defecto) en el modelo PCA: estadístico T^2 . Sin escalado previo. Escalado GS.....	87
Figura 6.16: Proyección de los primeros 30 experimentos (con defecto y sin defecto) en el modelo PCA: estadístico T^2 . Sin escalado previo. Escalado GS.....	87
Figura 6.17: Proyección de todos los experimentos (con defecto y sin defecto) en el modelo PCA: estadístico Q. Escalado previo. Escalado AS.....	88
Figura 6.18: Proyección de los primeros 30 experimentos (con defecto y sin defecto) en el modelo PCA: estadístico Q. Escalado previo. Escalado AS.....	89
Figura 6.19: Proyección de todos los experimentos (con defecto y sin defecto) en el modelo PCA: estadístico Q. Escalado previo. Escalado GS.....	89
Figura 6.20: Proyección de los primeros 30 experimentos (con defecto y sin defecto) en el modelo PCA: estadístico Q. Escalado previo. Escalado GS.....	90
Figura 6.21: Proyección de todos los experimentos (con defecto y sin defecto) en el modelo PCA: estadístico Q. Sin escalado previo. Escalado AS.....	91
Figura 6.22: Proyección de los primeros 30 experimentos (con defecto y sin defecto) en el modelo PCA: estadístico Q. Sin escalado previo. Escalado AS.....	91
Figura 6.23: Proyección de todos los experimentos (con defecto y sin defecto) en el modelo PCA: estadístico Q. Sin escalado previo. Escalado GS.....	92

Figura 6.24: Proyección de los primeros 30 experimentos (con defecto y sin defecto) en el modelo PCA: estadístico Q. Sin escalado previo. Escalado GS.....	92
Figura 6.25: Proyección de todos los experimentos (con defecto y sin defecto) en el modelo PCA: T^2 vs Q. Escalado previo. Escalado AS.....	94
Figura 6.26: Proyección de todos los experimentos (con defecto y sin defecto) en el modelo PCA: T^2 vs Q. Escalado previo. Escalado GS.....	94
Figura 6.27: Proyección de todos los experimentos (con defecto y sin defecto) en el modelo PCA: T^2 vs Q. Sin escalado previo. Escalado AS.....	95
Figura 6.28: Proyección de todos los experimentos (con defecto y sin defecto) en el modelo PCA: T^2 vs Q. Sin escalado previo. Escalado GS.....	96
Figura 7.1: Localización de defectos [14].....	99
Figura 7.2: Ejemplo de contribuciones por sensores.....	100

Bibliografía

- [1] Mujica, L.E.; Rodellar, J.; Vehí, J. A review of impact damage detection in structures using strain data. International Journal of COMADEM, 13, 3-18. 2010.
- [2] Sensor piezoeléctrico. (2018, 5 de Marzo). Wikipedia, The Free Encyclopedia. Disponible en <https://es.wikipedia.org/wiki/Sensor_piezoel%C3%A9ctrico>
- [3] Farrar, C.R.; Worden, K. An introduction to structural health monitoring. Philosophical Transactions of the Royal Society, 365, 303-315. 2006.
- [4] Dhakal, D.R.; Neupane, K.; Thapa, C.; Ramanjaneyulu, G.V. Different techniques of structural health monitoring. International Journal of Civil, Structural, Environmental and Infrastructure Engineering Research and Development (IJCSEIERD), 3 (2), 55-66. 2013.
- [5] Su, Z. and L. Ye. Identification of Damage Using Lamb Waves. Berlin: Springer, 2009
- [6] Ondas de Love. (2019, 18 de Enero). Wikipedia, The Free Encyclopedia. Disponible en: <https://es.wikipedia.org/wiki/Ondas_de_Love>
- [7] Tibaduiza, D.A.; Mujica, L.E.; Rodellar, J. Damage classification in structural health monitoring using principal component analysis and self-organizing maps. Structural control and health monitoring, 20:1303-1316. 2013.
- [8] Dell. Dell Latitude E6530 [en línea]. Disponible en: <https://www.dell.com/downloads/global/products/latit/dell_latitude_e6530_spec_sheet.pdf>

- [9] Krohn Hite Corporation. Wideband Amplifier Model 7600M/7602M [en línea]. Disponible en: <<http://www.krohn-hite.com/htm/amps/PDF/7600Data.pdf>>
- [10] Velleman. VTSS4 User manual [en línea]. Disponible en: <https://www.velleman.eu/downloads/1/vtss4gbnlfresd.pdf>
- [11] PCE-Iberica. Cámara térmica profesional PCE-TC 3 [en línea]. Disponible en: <<https://www.pce-iberica.es/medidor-detalles-tecnicos/instrumento-de-temperatura/camara-termica-s80hf.htm>>
- [12] Pico Technology. PicoScope 5000 A and B Series User's Guide [en línea]. Disponible en: <<https://www.picotech.com/download/manuals/ps5000ab.en-3.pdf>>
- [13] Ruiz, M.; Mujica, L.E.; Berjaga, X.; Rodellar, J. Partial least square/projection to latent structures (PLS) regression to estimate impact localization in structures. IOP Publishing: Smart Materials and Structures, 22. 2013.
- [14] Ruiz, M.; Mujica, L.E.; Sierra, J.; Pozo, F.; Rodellar, J. Multiway Principal Component Analysis Contributions for structural damage localization. 2016.
- [15] Mujica, L.E.; Rodellar, J.; Fernandez, A.; Güemes, A. Q-statistic and T2-statistic PCA-based measures for damage assessment in structures. Structural health monitoring: an international journal, 10, 539 - 553. 2010.
- [16] Kassambara, A. Practical Guide to Principal Components Methods in R. STHDA, 2017.
- [17] Radecki, R.; Staszewski, W.J.; Uhl, T. Impact of Changing Temperature on Lamb Wave Propagation for Damage Detection. Key Engineering Materials, 588, 140-148. 2013.

Detección de puntos calientes en estructuras utilizando ondas guiadas

- [18] Lamb waves. (2018, 7 de Mayo). Wikipedia, The Free Encyclopedia. Disponible en: <https://en.wikipedia.org/w/index.php?title=Lamb_waves&oldid=840082106>
- [19] Rayleigh wave. (2018, 20 de Marzo). Wikipedia, The Free Encyclopedia. Disponible en: <https://en.wikipedia.org/w/index.php?title=Rayleigh_wave&oldid=831336689>
- [20] National Instruments. NI PXI/PXI-5412 Specifications [en línea]. Disponible en: <<http://www.ni.com/pdf/manuals/371468e.pdf>>
- [21] National Instruments. PXI-5114 Specifications [en línea]. Disponible en: <<http://www.ni.com/pdf/manuals/374179g.pdf>>
- [22] National Instruments. NI PXI/PXIe-2529 Specifications [en línea]. Disponible en: <<http://www.ni.com/pdf/manuals/373649k.pdf>>
- [23] National Instruments. NI PXI-1033 User Manual and Specifications [en línea]. Disponible en: <<http://www.ni.com/pdf/manuals/371991c.pdf>>
- [24] Amprobe. IR608A Infrared Thermometer [en línea]. Disponible en: <<http://content.amprobe.com/DataSheets/IR608A%20Infrared%20Thermometer.pdf>>

Anexo A. Guía de manejo en R de los datos de los experimentos

A1. Preparación

Para manejar los datos mediante el software R se han utilizado varios paquetes pertenecientes al propio software: el paquete `ggplot2`, el cual nos permite elaborar gráficos a partir de un proceso de acumulación de capas; `reshape2`, que nos permite manipular y transformar los datos para darles la forma necesaria para su uso en un análisis posterior; y el paquete `dplyr`, que nos proporciona una forma ágil de manejar ficheros de datos en R.

Existen dos formas de instalar paquetes en R: la primera, a través de la pestaña de herramientas del software y seguidamente la opción instalar paquetes; la segunda, a través de nuestra consola o script utilizando la función “`install.packages()`”. Por ejemplo, si se desean instalar los paquetes citados anteriormente escribiremos:

```
install.packages("ggplot2")
install.packages("reshape2")
install.packages("dplyr")
```

Una vez instalados los paquetes necesarios, debemos borrar toda la memoria:

```
rm(list=ls())
```

Se cargan las librerías necesarias. Los paquetes solo se tienen que instalar una vez, pero se tienen que cargar siempre.

```
library(ggplot2)
library(reshape2)
library(dplyr)
```

Se carga el fichero de datos sobre el que trabajamos. Se puede cargar escribiendo el código que encontramos a continuación en nuestra consola o script, (cambiando la ruta del archivo, esta dependerá de cada usuario) o abriendo directamente el archivo con R.

```
load("C:/Users/Jesus/Desktop/tfg/experimento_final/Actuator_1.RData")
```

Detección de puntos calientes en estructuras utilizando ondas guiadas

El fichero de datos sobre el que se trabaja en el presente proyecto se ha obtenido a partir de los archivos pertenecientes a cada experimento realizado. Estos archivos de datos se han juntado en uno solo creando un data frame para que sea más fácil su manipulación en R.

Un data frame al igual que una matriz, representa una estructura rectangular de datos, es decir, almacenan datos en filas y columnas. La diferencia entre ambos, es que las matrices almacenan un único tipo de datos mientras que los data frames pueden contener distintos tipos.

El data frame generado en el presente proyecto, denominado Data, contiene 99000 observaciones (filas) y 154 variables (columnas).

Las filas están formadas por:

- 2750 muestras por cada señal (Inicialmente se tenían 6875 muestras por señal, pero al cortarlas han quedado 2750)
- 4 sensores
- Total de casos: 8 casos con defecto + baseline

Si multiplicamos cada una de las variables anteriores, obtenemos las 99000 observaciones.

Las variables son:

- Defecte: nos indica el tipo de caso al que pertenecen las muestras
- Actuator: número del actuador (siempre será 1, ya que solo se ha utilizado un actuador)
- Sensor: número del sensor (los 4 sensores utilizados están numerados del 2 al 5, ya que el número 1 pertenece al actuador)
- Time: Tiempo de la señal (debido al corte de las señales, no empieza en 0)
- Experimentos u observaciones O1-O150: 150 columnas, puesto que en el baseline se hicieron 150 experimentos mientras que para el resto solo 30. Las observaciones que correspondan al baseline tienen datos en las 150 columnas mientras que para el resto, las columnas del O31 al O150 no tienen datos (NA)

A2. Selección de experimentos

Si dentro de nuestro archivo, se desea seleccionar específicamente una serie de datos, acudiremos a la expresión lógica “subset”, que nos permite mantener los elementos que indiquemos.

A continuación encontramos una serie de ejemplos para seleccionar varios conjuntos de datos:

La siguiente línea nos permite seleccionar el experimento número 29 para todos los defectos y sensores:

```
new = subset(Data,select = c(Defecte,Sensor,029))
```

Si se quiere seleccionar, por ejemplo, sólo los experimentos del baseline, se ejecuta

```
baseline = subset(Data,Defecte=="Baseline")
```

Si queremos seleccionar los experimentos de un sensor específico, por ejemplo del sensor 2, ejecutamos:

```
Sensor2 = subset(Data,Sensor==2)
```

Si lo que se desea seleccionar específicamente los datos de un sensor en un determinado experimento, por ejemplo: los datos del sensor 2 en el experimento HS 0-2.

```
Sensor2_Defecte01 = subset(Data,Sensor==2 & Defecte=="HS 0-2")
```

A3. Representación gráfica

Para generar gráficos de los datos se deben seguir una serie de pasos:

1. Se deben seleccionar los datos deseados para graficar.
2. Se reorganizan los datos para poder representarlos en el mismo eje de tiempo.
3. Se crea el gráfico.

Por ejemplo, si deseamos graficar los datos de un determinado sensor para un determinado experimento, seleccionaremos los datos tal y como se ha explicado en el apartado anterior y añadiremos la función na.rm. Esta función elimina los valores NA (Not Available) siempre que se lo indiquemos “TRUE”.

Detección de puntos calientes en estructuras utilizando ondas guiadas

Ejemplo para el sensor dos y los datos del Baseline:

```
Sel_Data = subset(Data, Sensor==2 & Defecte=="Baseline", na.rm=TRUE)
```

Seguidamente, se reorganizan los datos. La función melt, perteneciente al paquete reshape2, es la responsable de reorganizar los datos.

```
mData<-melt (Sel_Data,id.vars=c("Time","Sensor","Defecte","Actuator"), na.rm=TRUE)
```

Para generar el gráfico ejecutamos:

```
ggplot(mData,aes(x=Time, y=value, colour=Defecte, group=variable)) + geom_path ()
```

La función aes, perteneciente al paquete ggplot2, describe cómo las variables de los datos se asignan a propiedades visuales (estéticas) de la geometría.

La función geom.path, también perteneciente al paquete ggplot2, nos permite conectar los datos en el orden en el que aparecen en el data.frame.

Si se desean graficar los datos de un sensor para más de un experimento, como por ejemplo los datos del sensor 2 para los defectos baseline y HS 0-2 (Figura 5.2), se hace de la siguiente forma:

```
Sel_Data = subset(Data,Sensor==2 & Defecte==c("Baseline","HS 0-2"), na.rm=TRUE)
mData      <-      melt(Sel_Data,id.vars=c("Time","Sensor","Defecte","Actuator"),na.rm=TRUE)
ggplot(mData,aes(x=Time, y=value, colour=Defecte, group=variable)) + geom_path()
```

Si se desea graficar los datos de un mismo sensor para todos los experimentos (Figura 5.3, Figura 5.4, Figura 5.5, Figura 5.6), se debe añadir una función. Esta función se denomina “facet_grid(vertical~horizontal)” (perteneciente al paquete ggplot2). Nos permite dividir los datos en una o dos variables tanto en dirección vertical como en horizontal. Puesto que nosotros queremos graficar todos los experimentos, llamaremos a la columna “Defecte” con la función facet_grid:

```
Sel_Data = subset(Data, Sensor==2, na.rm=TRUE)
mData <- melt(Sel_Data,id.vars=c("Time","Sensor","Defecte","Actuator"),na.rm=TRUE)
ggplot(mData,aes(x=Time, y=value, colour=Defecte, group=variable)) + geom_path() +
facet_grid(Defecte~.)
```

Anexo B. Guía de modelado de datos y realización de PCA

B1. Preparación

Para la realización de PCA a parte de los paquetes comentados en el anexo A, se han utilizado además otros paquetes: El paquete `mdatools`, el cual sirve para preprocesar, explorar y analizar conjuntos de datos multivariados; el paquete `data.table`, que nos permite organizar nuestros datos de forma rápida y sencilla; el paquete `tidyr`, también nos permite ordenar los datos de una forma sencilla y finalmente el paquete `ggalt`, este último paquete representa una ampliación para el paquete `ggplot2`, añadiendo varias formas y funciones para crear los gráficos.

Antes de empezar con el estudio, debemos instalar los paquetes.

```
install.packages("mdatools")
install.packages("ggplot2")
install.packages("reshape2")
install.packages("dplyr")
install.packages("data.table")
install.packages("tidyr")
install.packages("ggalt")
```

Se borra toda la memoria.

```
rm(list=ls())
```

Seguidamente cargamos las librerías necesarias para el modelado de los datos y la realización de PCA.

```
library(mdatools)
library(ggplot2)
library(reshape2)
library(dplyr)
library(data.table)
library(tidyr)
library(ggalt)
```

Antes de realizar PCA, debemos reorganizar nuestro `data.frame` y realizar el escalado de los datos. Para ello definimos una serie de funciones:

Detección de puntos calientes en estructuras utilizando ondas guiadas

1. Esta función sirve para cambiar los nombres de las columnas del data.frame. Se renombran de tal forma que nos indiquen el sensor y la muestra, por ejemplo: S1_3, indicando la muestra 3 del sensor 1.

```
cambiar_ColNames <-function(datos){  
a = c(rep("S1",2750),rep("S2",2750),rep("S3",2750),rep("S4",2750))  
b = as.character(rep(1:2750,4))  
c = data.frame(a,b)  
d = as.vector(unite(c,"c",c("a","b")))  
colnames(datos) = d$c  
datos  
}
```

2. La siguiente función agrega las variables defecto y observación al conjunto de datos.

```
agregar_ObsNames <-function(datos,obs_name){  
datos=cbind(Defecte=obs_name, Observation=1:nrow(datos),datos)  
datos  
}
```

3. Para realizar el escalado por groupscaling (GS), necesitaremos la siguiente función.

```
grp_escalado <-function(medias,desv_sen,datos){  
for (i in 1:nrow(datos)){  
vect_desv = c (rep (desv_sen [1],2750), rep(desv_sen[2],2750), rep(desv_sen[3],2750),  
rep(desv_sen[4],2750))  
datos_scaled[i,] = (datos[i,]-medias)/vect_desv  
}  
i  
datos_scaled  
}
```

4. La siguiente función realiza un escalado de los datos. Se ha utilizado para realizar una normalización previa al PCA, para eliminar el offset de algunas señales.

```
simple_escalado <-function(datos){  
datos_S1 = subset(datos,select = grepl("S1_", names(datos)))  
datos_S1_scaled = (datos_S1-apply(datos_S1,1,mean))/apply(datos_S1,1,sd)  
  
datos_S2 = subset(datos,select = grepl("S2_", names(datos)))  
datos_S2_scaled = (datos_S2-apply(datos_S2,1,mean))/apply(datos_S2,1,sd)  
  
datos_S3 = subset(datos,select = grepl("S3_", names(datos)))  
datos_S3_scaled = (datos_S3-apply(datos_S3,1,mean))/apply(datos_S3,1,sd)  
  
datos_S4 = subset(datos,select = grepl("S4_", names(datos)))
```

```
datos_S4_scaled = (datos_S4-apply(datos_S4,1,mean))/apply(datos_S4,1,sd)

data.frame(datos_S1_scaled,datos_S2_scaled,datos_S3_scaled,datos_S4_scaled)
}
```

B2. Preproceso de datos

Una vez definidas todas las funciones que se necesitarán más adelante, cargamos los datos y se preprocesan para realizar PCA.

Se separan los experimentos para realizar el entrenamiento, es decir, el modelo de PCA.

```
HS_baseline = Data %>%
subset(Defecte=="Baseline",select=c(01:0120)) %>% # Se seleccionan los datos
transpose() %>% # Se reorganizan para que una fila sea un experimento (todos los sensores)
cambiar_ColNames() %>% # Se asignan nuevo nombres a las columnas
simple_escalado() %>% # Se escala por separado cada sensor
agregar_ObsNames("Baseline")
```

La expresión “%>%” se utiliza para encadenar operaciones.

Se separan los experimentos para realizar la validación y se unen por filas.

```
val_baseline = Data %>% subset(Defecte=="Baseline",select=c(0121:0150)) %>% transpose() %>%
%
cambiar_ColNames() %>% simple_escalado() %>% agregar_ObsNames("No_HS")
val_HS01 = Data %>% subset(Defecte=="HS 0-1",select=c(02:030)) %>% transpose() %>%
cambiar_ColNames() %>% simple_escalado() %>% agregar_ObsNames("HS_0-1")
val_HS02 = Data %>% subset(Defecte=="HS 0-2",select=c(02:030)) %>% transpose() %>%
cambiar_ColNames() %>% simple_escalado() %>% agregar_ObsNames("HS_0-2")
val_HS03 = Data %>% subset(Defecte=="HS 0-3",select=c(02:030)) %>% transpose() %>%
cambiar_ColNames() %>% simple_escalado() %>% agregar_ObsNames("HS_0-3")
val_HS04 = Data %>% subset(Defecte=="HS 0-4",select=c(02:030)) %>% transpose() %>%
cambiar_ColNames() %>% simple_escalado() %>% agregar_ObsNames("HS_0-4")
val_HS12 = Data %>% subset(Defecte=="HS 1-2",select=c(02:030)) %>% transpose() %>%
cambiar_ColNames() %>% simple_escalado() %>% agregar_ObsNames("HS_1-2")
val_HS23 = Data %>% subset(Defecte=="HS 2-3",select=c(02:030)) %>% transpose() %>%
cambiar_ColNames() %>% simple_escalado() %>% agregar_ObsNames("HS_2-3")
val_HS34 = Data %>% subset(Defecte=="HS 3-4",select=c(02:030)) %>% transpose() %>%
cambiar_ColNames() %>% simple_escalado() %>% agregar_ObsNames("HS_3-4")
val_HS41 = Data %>% subset(Defecte=="HS 4-1",select=c(02:030)) %>% transpose() %>%
cambiar_ColNames() %>% simple_escalado() %>% agregar_ObsNames("HS_4-1")
```

Detección de puntos calientes en estructuras utilizando ondas guiadas

```
HS_validation =  
rbind(val_baseline, val_HS01, val_HS02, val_HS03, val_HS04, val_HS12, val_HS23, val_HS34, val_HS41  
)
```

Una vez unidos, eliminamos los experimentos por separado ya que ocupan memoria y ralentizan el programa.

```
rm(list=ls(pattern="val_"))
```

Llegados a este punto, se han obtenido dos data.frames: HS_baseline, que contiene los datos preprocesados para crear el modelo de PCA y HS_validation, que contiene los datos preprocesados para la validación.

B3. Creación del modelo y cálculo del PCA

Eliminamos las variables cuantitativas: Defecte y Observation

```
train=subset(HS_baseline, select = -c(Defecte, Observation))  
test=subset(HS_validation, select = -c(Defecte, Observation))
```

Aunque en el proyecto el análisis se ha realizado con autoscaling, a continuación se explicará como realizar el escalado de dos formas diferentes: Autoscaling (AS) y Groupscaling (GS).

1. Cálculo de PCA con autoscaling (AS)

```
model_PCA_AS = pca(train, 5, method='svd', scale=T) #El número 5 nos indica el número de  
componentes principales
```

Con la función method, seleccionamos el método de cálculo: Nipals o svd.

Se renombran los resultados, agregando las variables Defecte y Observation.

```
Baseline.Scores.AS=cbind(subset(HS_baseline, select=c(Defecte, Observation)), model_PCA_AS$ca  
lres$scores)
```

Se proyectan los nuevos datos:

```
Projected_Data_AS = predict(model_PCA_AS, test)
```

Se renombran los resultados de la proyección, agregando las variables Defecte y Observation.

```
Projected.Scores.AS=cbind(subset(HS_validation,select=c(Defecte,Observation)),Projected_Data_AS$scores)
```

2. Cálculo de PCA con groupscaling (GS)

La siguiente función es la encargada de realizar el escalado:

```
medias = colMeans(train)
desv_sen=c()
for (i in 1:4){
  k=i+1
  desv_sen[i] = sd(as.vector(data.matrix(subset(Data,Defecte=="Baseline" &
  Sensor==k,select=c(01:0120)))))
}
```

Se realiza el escalado:

```
train_grp_scaled = grp_escalado(medias,desv_sen,train)
test_grp_scaled = grp_escalado(medias,desv_sen,test)
```

Se aplica el PCA:

```
model_PCA_GS = pca(train_grp_scaled,5,method='svd',scale=F) #Se indica que no realice
autoscaling, scale=F (FALSE)
```

B4. Visualización de resultados de los scores

Se renombran los resultados, agregando las variables Defecte y Observation

```
Baseline.Scores.GS=cbind(subset(HS_baseline,select=c(Defecte,Observation)),model_PCA_GS$ca
lres$scores)
```

Se proyectan los nuevos datos

```
Projected_Data_GS = predict(model_PCA_GS,test_grp_scaled)
```

Se renombran los resultados de la proyección, agregando las variables Defecte y Observation

```
Projected.Scores.GS=cbind(subset(HS_validation,select=c(Defecte,Observation)),Projected_Data_GS$scores)
```

Visualización de los scores de los datos escalados con AS:

```
ScoresPlot.AS = ggplot(Baseline.Scores.AS, mapping = aes(x=`Comp 1`, y=`Comp 2`)) +
geom_point(aes(fill = Defecte), pch=25, size=3, colour="black") +
```

Detección de puntos calientes en estructuras utilizando ondas guiadas

```
geom_point(data=Projected.Scores.AS, aes(fill = Defecte), pch=21, size=3, colour="black")
plot(ScoresPlot.AS)
```

Visualización de los scores de los datos escalados con GS:

```
ScoresPlot.GS = ggplot(Baseline.Scores.GS, mapping = aes(x=`Comp 1`, y=`Comp 2`)) +
geom_point(aes(fill = Defecte), pch=25, size=3, colour="black") +
geom_point(data=Projected.Scores.GS, aes(fill = Defecte), pch=21, size=3, colour="black")
plot(ScoresPlot.GS)
```

La función `geom_point` se utiliza para diferenciar variables en gráficos. Con la función `aes` (aesthetics) se definen las propiedades visuales: *pch* para el tipo de forma (R le asigna un número a cada forma, habiendo 25 en total. Por ejemplo: el número 0 es un cuadrado blanco en su interior), *size* para el tamaño, *colour* para el color de contorno, *fill* para el color de relleno.

En nuestro caso, a la función *fill* no se le ha dado ningún color, sino que se ha llamado la variable Defecte. Esto genera automáticamente colores aleatorios para las diferentes variables existentes dentro de nuestra columna Defecte.

En caso de querer crear líneas envolventes en el gráfico de los scores, debemos seleccionar los datos de cada caso dentro del data.frame de los datos de los scores:

```
Scores.No_HS_circ <- Projected.Scores.AS[Projected.Scores.AS$Defecte == "No_HS", ]
Scores.HS_01_circ <- Projected.Scores.AS[Projected.Scores.AS$Defecte == "HS_0-1", ]
Scores.HS_02_circ <- Projected.Scores.AS[Projected.Scores.AS$Defecte == "HS_0-2", ]
Scores.HS_03_circ <- Projected.Scores.AS[Projected.Scores.AS$Defecte == "HS_0-3", ]
Scores.HS_04_circ <- Projected.Scores.AS[Projected.Scores.AS$Defecte == "HS_0-4", ]
Scores.HS_12_circ <- Projected.Scores.AS[Projected.Scores.AS$Defecte == "HS_1-2", ]
Scores.HS_23_circ <- Projected.Scores.AS[Projected.Scores.AS$Defecte == "HS_2-3", ]
Scores.HS_34_circ <- Projected.Scores.AS[Projected.Scores.AS$Defecte == "HS_3-4", ]
Scores.HS_41_circ <- Projected.Scores.AS[Projected.Scores.AS$Defecte == "HS_4-1", ]
```

Para visualizar el gráfico de los scores añadiendo las líneas envolventes se utiliza el siguiente código:

```
ScoresPlot.AS.lines = ggplot(Baseline.Scores.AS, mapping = aes(x=`Comp 1`, y=`Comp 2`)) +
geom_point(aes(fill = Defecte), pch=25, size=3, colour="black") +
geom_point(data=Projected.Scores.AS, aes(fill = Defecte), pch=21, size=3, colour="black")
+
geom_encircle(data=Scores.No_HS_circ, color="maroon2", size=1) +
```

```
geom_encircle(data=Scores.HS_01_circ, color="orange3", size=1) +
geom_encircle(data=Scores.HS_02_circ, color="yellow4", size=1) +
geom_encircle(data=Scores.HS_03_circ, color="limegreen", size=1) +
geom_encircle(data=Scores.HS_04_circ, color="mediumseagreen", size=1) +
geom_encircle(data=Scores.HS_12_circ, color="lightseagreen", size=1) +
geom_encircle(data=Scores.HS_23_circ, color="deepskyblue1", size=1) +
geom_encircle(data=Scores.HS_34_circ, color="mediumpurple2", size=1) +
geom_encircle(data=Scores.HS_41_circ, color="orchid1", size=1)+
geom_encircle(data=Baseline.Scores.AS, color="coral3", size=1)
plot(ScoresPlot.AS.lines)
```

La función `geom_encircle` es la encargada de generar las líneas envolventes. El problema es que al haber generado colores aleatorios con la función `fill`, si se desean tener las líneas de los mismos colores que los casos a envolver, se deben buscar y comparar con la gamma de colores de R para poder asignarlos con la función “color”.

En caso de querer generar las líneas envolventes para GS, tan solo debemos cambiar los nombres de los conjuntos de datos en el código visto anteriormente.

B5. Visualización de resultados para T2-statistic y Q-statistic

A continuación solo se muestra el código para los resultados en AS, si queremos los resultados con GS, tan solo se debe cambiar los nombres de los datos correspondientes.

Para representar los estadísticos T2 y Q, se deben sumar los valores de los componentes principales, formando así una sola columna.

Para el baseline:

```
T2statistic=rowSums(model_PCA_AS$calres$T2) #Suma por filas de los componentes principales
para T2 en el baseline
Qstatistic=rowSums(model_PCA_AS$calres$Q) #Suma por filas de los componentes principales
para Q en el baseline
```

Se añaden las variables Defecte y Observation a los resultados:

```
Baseline.T2.AS=cbind(subset(HS_baseline,select=c(Defecte,Observation)),model_PCA_AS$calres$
$T2,T2statistic)
Baseline.Q.AS=cbind(subset(HS_baseline,select=c(Defecte,Observation)),model_PCA_AS$calres$
Q,Qstatistic)
```


Detección de puntos calientes en estructuras utilizando ondas guiadas

Se juntan ambas columnas en un mismo data.frame para poder crear el gráfico T2 contra Q más adelante:

```
Baseline_T2_vs_Q=cbind(subset(HS_baseline,select=c(Defecte,Observation)),T2statistic,Qstatistic)
```

Realizamos el mismo proceso para los datos proyectados:

```
T2statistic=rowSums(Projected_Data_AS$T2)
Qstatistic=rowSums(Projected_Data_AS$Q)
#Se añaden las variables Defecte y Observation a los resultados
Projected.T2.AS=cbind(subset(HS_validation,select=c(Defecte,Observation)),Projected_Data_AS$T2,T2statistic)
Projected.Q.AS=cbind(subset(HS_validation,select=c(Defecte,Observation)),Projected_Data_AS$Q,Qstatistic)
#Se juntan T2 y Q para crear el gráfico T2 contra Q
Projected_T2_vs_Q=cbind(subset(HS_validation,select=c(Defecte,Observation)),T2statistic,Qstatistic)
```

Visualización de todos los experimentos para el estadístico T2:

```
T2Plot.AS = ggplot(Baseline.T2.AS, mapping = aes(x=`Observation`, y=`T2statistic`)) +
geom_point(aes(fill = Defecte), pch=25, size=3, colour="black") +
geom_point(data=Projected.T2.AS, aes(fill = Defecte), pch=21, size=3, colour="black")
plot(T2Plot.AS)
```

Al graficar todos los experimentos no se pueden apreciar muy bien los casos de los datos proyectados ya que tienen tan solo 30 experimentos frente a los 120 que tiene el modelo baseline. Por eso mismo, si se quiere realizar el gráfico mostrando tan solo 30 experimentos para cada caso, se debe ejecutar lo siguiente:

```
T2Plot.AS_30=ggplot(Baseline.T2.AS[1:30,],mapping=aes(x=`Observation`,y=`T2statistic`)) +
geom_point(aes(fill = Defecte), pch=25, size=3, colour="black") +
geom_point(data=Projected.T2.AS,aes(fill = Defecte), pch=21, size=3, colour="black")
plot(T2Plot.AS_30)
```

Visualización de los experimentos para el estadístico Q:

```
QPlot.AS = ggplot(Baseline.Q.AS, mapping = aes(x=`Observation`, y=`Qstatistic`)) +
geom_point(aes(fill = Defecte), pch=25, size=3, colour="black") +
geom_point(data=Projected.Q.AS, aes(fill = Defecte), pch=21, size=3, colour="black")
plot(QPlot.AS)
#Visualización del estadístico Q para los 30 primeros experimentos
```

```
QPlot.AS_30=ggplot(Baseline.Q.AS[1:30,],mapping=aes(x=`Observation`, y=`Qstatistic`)) +
geom_point(aes(fill = Defecte), pch=25, size=3, colour="black") +
geom_point(data=Projected.Q.AS, aes(fill = Defecte), pch=21, size=3, colour="black")
plot(QPlot.AS_30)
```

Creación del gráfico T2 contra Q:

```
T2vsQPlot.AS=ggplot(Baseline_T2_vs_Q, mapping = aes(x=`T2statistic`, y=`Qstatistic`)) +
geom_point(aes(fill = Defecte), pch=25, size=3, colour="black") +
geom_point(data=Projected_T2_vs_Q, aes(fill=Defecte), pch=21, size=3, colour="black")
plot(T2vsQPlot.AS)
```

Para realizar una línea envolvente para cada caso en el gráfico T2 contra Q, se debe realizar el mismo proceso que se ha explicado para los scores. Se seleccionan los datos a los que se les desea realizar la línea envolvente, obteniendo así un data.frame para cada caso:

```
T2Q.No_HS_circ <- Projected_T2_vs_Q[Projected_T2_vs_Q$Defecte == "No_HS",]
T2Q.HS_01_circ <- Projected_T2_vs_Q[Projected_T2_vs_Q$Defecte == "HS_0-1", ]
T2Q.HS_02_circ <- Projected_T2_vs_Q[Projected_T2_vs_Q$Defecte == "HS_0-2", ]
T2Q.HS_03_circ <- Projected_T2_vs_Q[Projected_T2_vs_Q$Defecte == "HS_0-3", ]
T2Q.HS_04_circ <- Projected_T2_vs_Q[Projected_T2_vs_Q$Defecte == "HS_0-4", ]
T2Q.HS_12_circ <- Projected_T2_vs_Q[Projected_T2_vs_Q$Defecte == "HS_1-2", ]
T2Q.HS_23_circ <- Projected_T2_vs_Q[Projected_T2_vs_Q$Defecte == "HS_2-3", ]
T2Q.HS_34_circ <- Projected_T2_vs_Q[Projected_T2_vs_Q$Defecte == "HS_3-4", ]
T2Q.HS_41_circ <- Projected_T2_vs_Q[Projected_T2_vs_Q$Defecte == "HS_4-1", ]
```

Para generar el gráfico con las líneas envolventes se ejecuta:

```
plot(ggplot(Baseline_T2_vs_Q, mapping = aes(x=`T2statistic`, y=`Qstatistic`)) +
geom_point(aes(fill = Defecte), pch=25, size=3, colour="black") +
geom_point(data=Projected_T2_vs_Q, aes(fill = Defecte), pch=21, size=3, colour="black")+
geom_encircle(data=T2Q.No_HS_circ, color="maroon2", size=1) +
geom_encircle(data=T2Q.HS_01_circ, color="orange3", size=1) +
geom_encircle(data=T2Q.HS_02_circ, color="yellow4", size=1) +
geom_encircle(data=T2Q.HS_03_circ, color="limegreen", size=1) +
geom_encircle(data=T2Q.HS_04_circ, color="mediumseagreen", size=1) +
geom_encircle(data=T2Q.HS_12_circ, color="lightseagreen", size=1) +
geom_encircle(data=T2Q.HS_23_circ, color="deepskyblue1", size=1) +
geom_encircle(data=T2Q.HS_34_circ, color="mediumpurple2", size=1) +
geom_encircle(data=T2Q.HS_41_circ, color="orchid1", size=1)+
geom_encircle(data=Baseline_T2_vs_Q, color="coral3", size=1))
```

Anexo C. Introducción al cálculo de contribuciones con R

En el anexo C supondremos que seguimos trabajando con los datos y los paquetes comentados anteriormente en los anexos A y B.

C1. Cálculo de Q por residuals y por loadings

Una vez se ha calculado PCA tanto para GS como para AS, si queremos calcular contribuciones se deben reorganizar los datos de la siguiente forma:

```
Results = cbind(subset(HS_baseline,select =  
c(Defecte,Observation)),model_PCA$calres$scores)  
  
Results = rbind(Results, cbind(subset(HS_validation,select =  
c(Defecte,Observation)),Projected_Data$scores))
```

En el data frame “Results” se han organizado los datos del training más los del testing.

Se agrega a “Results” la suma de Q por cada PC:

```
Results=cbind(Results,Q.stat = c(rowSums(model_PCA$calres$Q), rowSums(Projected_Data$Q)))
```

Se calcula Q a partir de los residuals, para ello utilizamos la siguiente función:

```
Q.resi=vector()  
for (i in 1:dim(model_PCA$calres$residuals)[1]){ # Training  
  e = model_PCA$calres$residuals[i,]  
  Q.resi = c(Q.resi,t(e)%*%e)  
}  
  
for (i in 1:dim(Projected_Data$residuals)[1]){ # testing  
  e = Projected_Data$residuals[i,]  
  Q.resi = c(Q.resi,t(e)%*%e)  
}  
  
Results=cbind(Results,Q.resi)
```

También podemos calcular Q a partir de los loadings, el resultado debe ser el mismo:

```

Q.load=vector()
P = model_PCA$loadings
IP = diag(1,nrow = dim(P)[1])-P%*%t(P)

train = as.matrix(subset(HS_baseline,select = -c(Defecte,Observation)))
for (i in 1:dim(model_PCA$calres$residuals)[1]){ # Training
  print(i)
  x = as.vector((train[i,]-model_PCA$center)/model_PCA$scale)
  Q.load = c(Q.load,x%*%IP%*%x)
}

test = as.matrix(subset(HS_validation,select = -c(Defecte,Observation)))
for (i in 1:dim(Projected_Data$residuals)[1]){ # Testing
  print(i)
  x = as.vector((test[i,]-model_PCA$center)/model_PCA$scale)
  Q.load = c(Q.load,x%*%IP%*%x)
}
Results=cbind(Results,Q.load)

```

C2. Cálculo y representación gráfica de las contribuciones de Q

El cálculo para las contribuciones de Q se ha realizado con la Q calculada a partir de los residuals.

Primero se añaden las variables Defecte y Observation a los resultados de Q anteriormente calculados.

```

Contr.Q.res = cbind(subset(HS_baseline,select =
c(Defecte,Observation)),model_PCA$calres$residuals^2)

Contr.Q.res = rbind(Contr.Q.res,cbind(subset(HS_validation,select =
c(Defecte,Observation)),Projected_Data$residuals^2))

```

Seguidamente se calcula su contribución, para ello se utiliza el siguiente código:

```

Contr.Q.org = data.frame()
def = levels(Contr.Q.res$Defecte)
for (i in 1:length(def)){
  sel_data = subset(Contr.Q.res,Defecte==def[i]) %>%
    melt(id.vars=c("Defecte","Observation"),na.rm=TRUE) %>%
    spread(Observation,value)
  Contr.Q.org = bind_rows(Contr.Q.org,sel_data)
}
Contr.Qres.org = select(Contr.Q.org, -variable)
Sensor=rep(c(rep(1,2750),rep(2,2750),rep(3,2750),rep(4,2750)),10)

```

Detección de puntos calientes en estructuras utilizando ondas guiadas

```
Time=rep(seq(1:11000),10)
Contr.Qres.org = cbind(Sensor,Time,Contr.Qres.org)
```

Por último, se grafican los resultados:

```
Sel_Data = subset(Contr.Qres.org, select=Sensor:4)
mData <- melt(Sel_Data,id.vars=c("Defecte","Sensor","Time"),na.rm=TRUE) #
ggplot(mData,aes(x=Time, y=value, colour=Defecte, group=variable)) + geom_path() +
facet_grid(Defecte~.)
```